

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: September 27, 2004, 07:46:34 ; Search time 22285 Seconds

(without alignments)  
11591.855 Million cell updates/sec

Title: US-10-030-294-2

Perfect score: 5960

Sequence: 1 cagcgttcagagatcctcgtc.....tttgcgtcatatttcgac 5960

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenBank:\*

1: gb\_ba:\*

2: gb\_hcg:\*

3: gb\_in:\*

4: gb\_ov:\*

5: gb\_ov:\*

6: gb\_ov:\*

7: gb\_ov:\*

8: gb\_ov:\*

9: gb\_ov:\*

10: gb\_ov:\*

11: gb\_ov:\*

12: gb\_ov:\*

13: gb\_ov:\*

14: gb\_ov:\*

15: gb\_ov:\*

16: gb\_ov:\*

17: gb\_ov:\*

18: gb\_ov:\*

19: gb\_ov:\*

20: gb\_ov:\*

21: gb\_ov:\*

22: gb\_ov:\*

23: gb\_ov:\*

24: gb\_ov:\*

25: gb\_ov:\*

26: gb\_ov:\*

27: gb\_ov:\*

28: gb\_ov:\*

29: gb\_ov:\*

30: gb\_ov:\*

31: gb\_ov:\*

32: gb\_ov:\*

33: gb\_ov:\*

34: gb\_ov:\*

35: gb\_ov:\*

36: gb\_ov:\*

37: gb\_ov:\*

38: gb\_ov:\*

39: gb\_ov:\*

40: gb\_ov:\*

41: gb\_ov:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5960	100.0	5960	6	BD093090 Gene enco
2	5919.6	99.3	152646	6	BD093090 Gene enco
3	5681.6	95.3	13940	6	BD093090 Gene enco
4	5676	95.2	5676	6	BD093090 Gene enco
5	5676	95.2	5676	6	BD093090 Gene enco
6	5676	95.2	5676	6	BD093090 Gene enco
7	843.2	14.1	160574	2	AC135489 Rattus no
8	843.2	14.1	284005	2	AC137884 Rattus no
9	835.8	14.0	204333	2	AC126304 Rattus no
10	835.8	14.0	214853	10	AC113020 Rattus no
11	499	8.4	835	9	AF3596351 Mus muscu
12	470.4	7.9	531	9	AF3596351 Mus muscu
13	402.2	6.7	56268	2	AF533892 Mus muscu
14	273.4	4.6	208158	10	AF533892 Mus muscu
15	212.4	3.6	159249	9	AC103882 Homo sapi
16	205	3.4	60323	2	AC087334 Homo sapi
17	204.4	3.4	173456	9	AL158198 Homo sapi
18	201.2	3.4	150846	2	AC036173 Homo sapi
19	201.2	3.4	173178	2	AC135613 Homo sapi
20	200	3.4	159808	2	AL591182 Homo sapi
21	196.6	3.3	123377	2	AC118135 Homo sapi
22	196.6	3.3	181442	2	AC134875 Homo sapi
23	196.6	3.3	202827	9	AC016907 Homo sapi
24	196.2	3.3	219122	9	AC080038 Homo sapi
25	196	3.3	148911	2	AL353142 Homo sapi
26	195	3.3	273403	2	AC011498 Homo sapi
27	195	3.3	47028	9	AC117488 Homo sapi
28	194	3.3	163246	2	AC068401 Homo sapi
29	194	3.3	174428	9	AC068401 Homo sapi
30	193.4	3.2	118384	9	AF389851 Homo sapi
31	193.2	3.2	196467	9	AC022352 Homo sapi
32	193	3.2	139218	2	AC019289 Homo sapi
33	193	3.2	151880	9	AP005716 Homo sapi
34	193	3.2	166093	2	AC106013 Homo sapi
35	193	3.2	175931	2	AC026618 Homo sapi
36	193	3.2	180776	2	AC032043 Homo sapi
37	193	3.2	183336	9	AC108688 Homo sapi
38	193	3.2	229641	2	AC092893 Homo sapi
39	191.8	3.2	179305	2	AC144411 Homo sapi
40	191.2	3.2	131757	9	AC004540 Homo sapi
41	190.6	3.2	78340	9	AL359473 Homo sapi
42	190.6	3.2	161416	2	AC026857 Homo sapi
43	190.6	3.2	218074	2	AC002044 Homo sapi
44	190.4	3.2	157136	2	AC068251 Homo sapi
45	190.4	3.2	165240	9	AP002795 Homo sapi

## ALIGNMENTS

RESULT 1

BD093090 5960 bp DNA linear PAT 27-AUG-2002

LOCUS BD093090

DEFINITION Gene encoding promoter region of tumor suppressor gene p51.

ACCESSION BD093090.1 GI:22638678

VERSION WO 0100818-A/2.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 5960)

AUTHORS Sakai,T., Kagaya,S., Sato,T., Suenaga,Y. and Fujii,H.

TITLE Gene encoding promoter region of tumor suppressor gene p51

JOURNAL Patent: WO 0100818-A 2 04-JAN-2001.

Pred. No. is the number of results predicted by chance to have a



2 30 11:51:28 2004

us-10-030-294-2.std.rge

1741 ATAAATTTAGCTCTGAGTGGCTTACCCATGTAAGTGGCTACCAATCAGAC 1800  
QY 1801 GGTGAAATTTAGTATTTCTTATTAAGAAAGCTGATTTGAAAAATGTTCTAG 1860  
Db 1801 GGTGAAATTTAGTATTTCTTATTAAGAAAGCTGATTTGAAAAATGTTCTAG 1860  
QY 1861 AAAATTTACATTAATTTAAATTAAGAAAGCTGATTTGAAAAATGTTCTAG 1920  
Db 1861 AAAATTTACATTAATTTAAATTAAGAAAGCTGATTTGAAAAATGTTCTAG 1920  
QY 1921 TGAAGGAGATGATTTGAAAGGATTTGAAAGGATTTGAAAGGATTTGAAAG 1980  
Db 1921 TGAAGGAGATGATTTGAAAGGATTTGAAAGGATTTGAAAGGATTTGAAAG 1980  
QY 1981 ATCTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
Db 1981 ATCTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040  
QY 2041 TGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
Db 2041 TGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
QY 2101 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
Db 2101 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
QY 2161 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220  
Db 2161 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2220  
QY 2221 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
Db 2221 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
QY 2281 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
Db 2281 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
QY 2341 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
Db 2341 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
QY 2401 GAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460  
Db 2401 GAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460  
QY 2461 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2520  
Db 2461 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2520  
QY 2521 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
Db 2521 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
QY 2581 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
Db 2581 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
QY 2641 GAGTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
Db 2641 GAGTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
QY 2701 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760  
Db 2701 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760  
QY 2761 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
Db 2761 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
QY 2821 GTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880  
Db 2821 GTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880

QY 2881 CTGACATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940  
Db 2881 CTGACATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940  
QY 2941 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000  
Db 2941 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000  
QY 3001 TTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
Db 3001 TTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
QY 3061 CCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120  
Db 3061 CCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120  
QY 3121 TTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180  
Db 3121 TTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180  
QY 3181 ACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
Db 3181 ACTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
QY 3241 TTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300  
Db 3241 TTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300  
QY 3301 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360  
Db 3301 CATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360  
QY 3361 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420  
Db 3361 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420  
QY 3421 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480  
Db 3421 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480  
QY 3481 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540  
Db 3481 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540  
QY 3541 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600  
Db 3541 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600  
QY 3601 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660  
Db 3601 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660  
QY 3661 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720  
Db 3661 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720  
QY 3721 TTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3780  
Db 3721 TTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3780  
QY 3781 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3840  
Db 3781 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3840  
QY 3841 AGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900  
Db 3841 AGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900  
QY 3901 GCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3960  
Db 3901 GCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3960

Thu Sep 30 11:51:28 2004

us-10-030-294-2.Std.rge

Page 4

QY 3961 ATTATTAAGGATGACCTTGGGCTGGTGCACAGGAAATGATTTTATTTATTCAGAA 4020  
| | | | |  
DB 3961 ATTATTAAGGATGACCTTGGGCTGGTGCACAGGAAATGATTTTATTTATTCAGAA 4020  
QY 4021 CCTTCTATTAGGTCATCTATATTTGCTTAATGCGAGGAGAAAGCCAACTCTTTAACT 4080  
| | | | |  
DB 4021 CCTTCTATTAGGTCATCTATATTTGCTTAATGCGAGGAGAAAGCCAACTCTTTAACT 4080  
QY 4081 GCAATTAACAAATCTATATTAATTAATGATTAGCAATCTTCCCTTTAAGTTTCAATTTG 4140  
| | | | |  
DB 4081 GCAATTAACAAATCTATATTAATTAATGATTAGCAATCTTCCCTTTAAGTTTCAATTTG 4140  
QY 4141 TCGACAGAGCTGTTGATTTGGCTGGGCTGAGGCGGCTGTTGGATTGCAAAAT 4200  
| | | | |  
DB 4141 TCGACAGAGCTGTTGATTTGGCTGGGCTGAGGCGGCTGTTGGATTGCAAAAT 4200  
QY 4201 CACAGATTTAGCGGCTCTCGGCTAAGTAAGAGAGAAATGTCAGATTTAATAGCT 4260  
| | | | |  
DB 4201 CACAGATTTAGCGGCTCTCGGCTAAGTAAGAGAGAAATGTCAGATTTAATAGCT 4260  
QY 4261 TCTCCCTTCATCTGGCTGAGACACAAATTAATTTTATGAAACATTTTGAGT 4320  
| | | | |  
DB 4261 TCTCCCTTCATCTGGCTGAGACACAAATTAATTTTATGAAACATTTTGAGT 4320  
QY 4321 TAGATTTACTACAGGGAATGTCATTTCTGGAAGGCTTAAGATGTTCTCAAC 4380  
| | | | |  
DB 4321 TAGATTTACTACAGGGAATGTCATTTCTGGAAGGCTTAAGATGTTCTCAAC 4380  
QY 4381 TTGATTTACTACAGGGAATGTCATTTCTGGAAGGCTTAAGATGTTCTCAAC 4440  
| | | | |  
DB 4381 TTGATTTACTACAGGGAATGTCATTTCTGGAAGGCTTAAGATGTTCTCAAC 4440  
QY 4441 AGATGTAACCTCACGATTTAGTACCGTTAGATACAGAGAGCTTTTTCCTCCG 4500  
| | | | |  
DB 4441 AGATGTAACCTCACGATTTAGTACCGTTAGATACAGAGAGCTTTTTCCTCCG 4500  
QY 4501 TTGAGATTTACTACAGGGAATGTCATTTCTGGAAGGCTTAAGATGTTCTCAAC 4560  
| | | | |  
DB 4501 TTGAGATTTACTACAGGGAATGTCATTTCTGGAAGGCTTAAGATGTTCTCAAC 4560  
QY 4561 ACGAGATTAAGTTTACGAGCCATATTCAGAAAGGAATTAATTTTGTGTAGACT 4620  
| | | | |  
DB 4561 ACGAGATTAAGTTTACGAGCCATATTCAGAAAGGAATTAATTTTGTGTAGACT 4620  
QY 4621 TCGGATTTACTACAGGGAATGTCATTTCTGGAAGGCTTAAGATGTTCTCAAC 4680  
| | | | |  
DB 4621 TCGGATTTACTACAGGGAATGTCATTTCTGGAAGGCTTAAGATGTTCTCAAC 4680  
QY 4681 GTCAAGTCAGAGCAAAACCAAAACGCAAAACGTAAGCATTAAGATGAGT 4740  
| | | | |  
DB 4681 GTCAAGTCAGAGCAAAACCAAAACGCAAAACGTAAGCATTAAGATGAGT 4740  
QY 4741 CCGACTGAGATTTAAATTAATGATTAATTTTATTAACAGGCAATTTGAAATTA 4800  
| | | | |  
DB 4741 CCGACTGAGATTTAAATTAATGATTAATTTTATTAACAGGCAATTTGAAATTA 4800  
QY 4801 GTGACATTCAGAAATTAATTAATTAATTTTATTAACAGGCAATTTGAAATTA 4860  
| | | | |  
DB 4801 GTGACATTCAGAAATTAATTAATTAATTTTATTAACAGGCAATTTGAAATTA 4860  
QY 4861 TTAATTAATTAATTAATTAATTAATTTTATTAACAGGCAATTTGAAATTA 4920  
| | | | |  
DB 4861 TTAATTAATTAATTAATTAATTAATTTTATTAACAGGCAATTTGAAATTA 4920  
QY 4921 AGGTATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4980  
| | | | |  
DB 4921 AGGTATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4980  
QY 4981 CTTTCTCTTACGAG 5040  
| | | | |  
DB 4981 CTTTCTCTTACGAG 5040  
QY 5041 GTCTTTTCTATTTTGAATTTGTGACCAACAAGCGGCTTGCTGAAGAGAACTGAAGG 5100  
| | | | |

DB 5041 GTCTTTTCTATTTTGAATTTGTGACCAACAAGCGGCTTGCTGAAGAGAACTGAAGG 5100  
QY 5101 CCGGAG 5160  
| | | | |  
DB 5101 CCGGAG 5160  
QY 5161 ACAAACATTTTACCCCAAGAAATAGTCAAGAAATCTCAATTAATCAACAGTATCCGA 5220  
| | | | |  
DB 5161 ACAAACATTTTACCCCAAGAAATAGTCAAGAAATCTCAATTAATCAACAGTATCCGA 5220  
QY 5221 TACAG 5280  
| | | | |  
DB 5221 TACAG 5280  
QY 5281 AAGATCCAGGCTGCAATTAATTAATCTGATGATCAATTCAGAGATCAATCAAGACA 5340  
| | | | |  
DB 5281 AAGATCCAGGCTGCAATTAATTAATCTGATGATCAATTCAGAGATCAATCAAGACA 5340  
QY 5341 GAGATCAAGATTCAG 5400  
| | | | |  
DB 5341 GAGATCAAGATTCAG 5400  
QY 5401 TACGTCAG 5460  
| | | | |  
DB 5401 TACGTCAG 5460  
QY 5461 GCTGTCAG 5520  
| | | | |  
DB 5461 GCTGTCAG 5520  
QY 5521 TGAAGAAAG 5580  
| | | | |  
DB 5521 TGAAGAAAG 5580  
QY 5581 TGTCTGATGAGATTTGACCCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5640  
| | | | |  
DB 5581 TGTCTGATGAGATTTGACCCATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5640  
QY 5641 AGGTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5700  
| | | | |  
DB 5641 AGGTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5700  
QY 5701 GGAATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5760  
| | | | |  
DB 5701 GGAATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5760  
QY 5761 CCGAGGTCAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5820  
| | | | |  
DB 5761 CCGAGGTCAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5820  
QY 5821 ATTATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5880  
| | | | |  
DB 5821 ATTATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5880  
QY 5881 TATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5940  
| | | | |  
DB 5881 TATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5940  
QY 5941 TTTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5960  
| | | | |  
DB 5941 TTTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 5960

RESULT 2  
AC078809/c 162646 bp DNA linear PRI 25-JAN-2003  
LOCUS AC078809 Homo sapiens 3 BAC Rpl1-60063 (Rosewell Park Cancer Institute Human  
DEFINITION BAC Library) complete sequence.  
ACCESSION AC078809  
VERSION AC078809.9 GI:1903390  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)





## BEST AVAILABLE COPY

```

repeat_region 7276. -7892
/rpt_family="LIMC3"
complement(8145. -8365)
repeat_region /rpt_family="MIR"
9447. -9479
/rpt_family="AT-rich"
9648. -9780
/rpt_family="CA-rich"
10440. -10660
/rpt_family="MSR20"
complement(10703. -11014)
repeat_region /rpt_family="AluSk"
11431. -11683
/rpt_family="AluSk"
11431. -11683
/standard_name="79189"
complement(11911. -12013)
repeat_region /rpt_family="MIR"
12654. -12964
/rpt_family="AluSk"
13022. -13142
/rpt_family="MIR"
15041. -15357
/rpt_family="AluSk"
15456. -15480
/rpt_family="TTTA"
complement(16012. -16151)
repeat_region /rpt_family="MIR"
17138. -17449
/rpt_family="AluSp"
17765. -17895
/rpt_family="L2"
18694. -18860
/rpt_family="MIR"
18907. -19141
/rpt_family="MIR"
19474. -19515
/rpt_family="(TA)n"
complement(19940. -20118)
repeat_region /rpt_family="AluSk"
20726. -21393
/rpt_family="LIM3e"
complement(21402. -21704)
repeat_region /rpt_family="L1"
21710. -23051
/rpt_family="LIM2"
23048. -24606
/rpt_family="LIM1"
24611. -24765

```

Query Matched	99.3%:	Score 5919.6:	DB 9:	Length 162646:
Best Local Similarity	99.8%:	Prod. No. 0:		
Matches 5950:	Conservative 0:	Mismatches 9:	Indels 5:	Gaps 2:
Qy	1	CAGCTGTTGAGGATGTCGAAAAAGAACCCACACATTCCTCTGGAACATCGGGTGT	60	
Db	76260	CAGCTGTTGAGGATGTCGAAAAAGAACCCACACATTCCTCTGGAACATCGGGTGT	76209	
Qy	61	GACTTGGAGGGTATCAGCTTGTCTTAAABAACTGCGCAACTCTTTCCTGCCCAAT	120	
Db	76200	GACTTGGAGGGTATCAGCTTGTCTTAAABAACTGCGCAACTCTTTCCTGCCCAAT	76141	
Qy	121	TGGCCTCTTTCCTTCGATGCCCTCTTTCCTTGGACACTCCCTTAAGCAATCTTCTG	180	
Db	76140	TGGCCTCTTTCCTTCGATGCCCTCTTTCCTTGGACACTCCCTTAAGCAATCTTCTG	76081	
Qy	181	ACATTACTTAAGTTAAATGTTTAAATGTAATTCGTAACCTGAGAGGAAATGGAG	240	
Db	76080	ACATTACTTAAGTTAAATGTTTAAATGTAATTCGTAACCTGAGAGGAAATGGAG	76021	
Qy	241	GTCGAATCGAAGAAAGCATGGCTAAGGTTGCATGCACTTGCTTTTTCATGAATTA	300	
Db	76020	GTCGAATCGAAGAAAGCATGGCTAAGGTTGCATGCACTTGCTTTTTCATGAATTA	75961	
Qy	301	AGCATTCGAAATACATTCAGTTTACTTAAGTTCTAGGCCCAAGTTTACTCTAATGAT	360	

Dp	75960	AGTCATTCGAATACCATTCAGTTTACTTAAAGTCTTACGCCACAGCTTATCTCTATATGAT	75901
Qy	361	GTCAAGCTGTAGCAAAATTTAGCTCCAAAAGTTGAAAGATTACAGCATCTCTCCATCA	420
Dp	75900	GTACAGCTGTACCAAAATTTAGGCTCCAAAGTTGAAAGATTACAGCATCTCTCCCATCA	75841
Qy	421	CAGAACTTTGGCTTCCATTTAACTAAATTAGAATTTGTGTGGTTGAAGCTCAGCTATGTA	480
Dp	75840	CAGAACTTTGGCTTCCATTTAACTAAATTAGAATTTGTGTGGTTGAAGCTCAGCTATGTA	75781
Qy	481	CAGAAAAGTCAATCAATTTAAAAATCACAAACCTCAGTCTCTCAATTTGAGCAATTAAG	540
Dp	75780	CAGAAAAGTCAATCAATTTAAAAATCACAAACCTCAGTCTCTCAATTTGAGCAATTAAG	75721
Qy	541	TGGTGAATTTAATCTCAACAAGCTCTCTCCCTGTAAGATTTCTTCCCTGCTCTCTCACTATA	600
Dp	75720	TGGTGAATTTAATCTCAACAAGCTCTCTCTCCCTGTAAGATTTCTTCCCTGCTCTCTCACTATA	75661
Qy	601	AATGCAGATGACCTGGAAAGGCTGAGACCTGAGGTTCAATTACCTTGACCAAAAGAAAT	660
Dp	75660	AATGCAGATGACCTGGAAAGGCTGAGACCTGAGGTTCAATTACCTTGACCAAAAGAAAT	75601
Qy	661	TCAGTTTCTGTATCTCATATGTACAAGGCTGCAGAGCTTACAGAACTACGAAGATTCAT	720
Dp	75600	TCAGTTTCTGTATCTCATATGTACAAGGCTGCAGAGCTTACAGAACTACGAAGATTCAT	75541
Qy	721	CTGCTTTAAGCCTCTGTGGGAGCATCTATGTTTTCACATGGCCCTTACTATTTGCTCTC	780
Dp	75540	CTGCTTTAAGCCTCTGTGGGAGCATCTATGTTTTCACATGGCCCTTACTATTTGCTCTC	75481
Qy	781	TTCTCTGCTTAAAGAACTTTTATTTTCTTCTGAAAACCTCTCTGCTCATGTCATGTTAGGG	840
Dp	75480	TTCTCTGCTTAAAGAACTTTTATTTTCTTCTGAAAACCTCTCTGCTCATGTCATGTTAGGG	75421
Qy	841	CCATCAGTCCACATGATTCAGGCTCTCTCTGAGCAAAACAATGGCATCTTCTTTGGGAAT	900
Dp	75420	CCATCAGTCCACATGATTCAGGCTCTCTCTGAGCAAAACAATGGCATCTTCTTTGGGAAT	75361
Qy	901	TCAATCTTAAGCTGAATAGCTGAAGTTTCAAAAAAGCTGTAATTCGACTTAAGGCTAC	960
Dp	75360	TCAATCTTAAGCTGAATAGCTGAAGTTTCAAAAAAGCTGTAATTCGACTTAAGGCTAC	75301
Qy	961	AGTGGCTTGAAGATGACATCTGCATTTCTTCAAGTCTGAACTTATTAATTTAT	1020
Dp	75300	AGTGGCTTGAAGATGACATCTGCATTTCTTCAAGTCTGAACTTATTAATTTAT	75241
Qy	1021	CCTGGTAAAGCCCTTCTTGAGATGTGTCTTTTTCACATCTGCTCTTAATAGTCTGT	1080
Dp	75240	CCTGGTAAAGCCCTTCTTGAGATGTGTCTTTTTCACATCTGCTCTTAATAGTCTGT	75181
Qy	1081	GAAATTTCAATTTCTTTTCAATTCATTTCAATGCT---TTTGTTGTTGTTTGTGTGTT	1136
Dp	75180	GAAATTTCAATTTCTTTTCAATTCATTTCAATGCT---TTTGTTGTTGTTTGTGTGTT	75121
Qy	1137	TTTGGCTTTAGAGGACAAATCAATTTCTGTGTGTTTATACCAAGAAATCCATATTA	1196
Dp	75120	TTTGGCTTTAGAGGACAAATCAATTTCTGTGTGTTTATACCAAGAAATCCATATTA	75061
Qy	1197	TACATTCCTTCCCTTTTAAAAATTAAGATTCGAAGGCTCAAAAGAGATGAGCTACCTGCT	1256
Dp	75060	TACATTCCTTCCCTTTTAAAAATTAAGATTCGAAGGCTCAAAAGAGATGAGCTACCTGCT	75001
Qy	1257	GAGGCTCTGGAGTAAGTTAGTACAGAGCTGTGTACTAAACCAGTTAGGCAATGCTTT	1316
Dp	75000	GAGGCTCTGGAGTAAGTTAGTACAGAGCTGTGTACTAAACCAGTTAGGCAATGCTTT	74941
Qy	1317	ACACACAATTTGCTCTCTCTTCAAGATTATAGCAAGCTTGTGAAAGAAAGAGCTACTATT	1376
Dp	74940	ACACACAATTTGCTCTCTCTTCAAGATTATAGCAAGCTTGTGAAAGAAAGAGCTACTATT	74881
Qy	1377	TTGCCAAGAAGCTCAGAGGACCAAGAACAGTTCTGGAGAAATCTATGATTTAACTCTT	1436

Db 74880 TTGCCAAGACCTCAGAGGACCAAGAACAGTTCTGGATATGTGATGATGAACCTT 74821  
Qy 1437 AAAAGTTTGTGACCTCTGGCCATATTTGTATCTTAAGCCAGATTTTCAATTTAT 1496  
Db 74820 AAAAGTTTGTGACCTCTGGCCATATTTGTATCTTAAGCCAGATTTTCAATTTAT 74761  
Qy 1497 AGCTAAACAAACAAAGAGATCCAGAGTTGAGAGCTATPATTAAGTGAATTAAGT 1556  
Db 74760 AGCTAAACAAACAAAGAGATCCAGAGTTGAGAGCTATPATTAAGTGAATTAAGT 74701  
Qy 1557 ATAGCTGACAGATGATATATCTGAGAACCAATGAGTCAAGAGAGAAAAATCC 1616  
Db 74700 ATAGCTGACAGATGATATATCTGAGAACCAATGAGTCAAGAGAGAAAAATCC 74641  
Qy 1617 AGACAGAGAGAAATGATACCAATATGCTGATTTATGTATGATTTCTAGCATTTATGT 1676  
Db 74640 AGACAGAGAGAAATGATACCAATATGCTGATTTATGTATGATTTCTAGCATTTATGT 74581  
Qy 1677 CCAACATGATGACAGATGACCTCAATGCTAATCAATTAAGTGAATTAAGT 1736  
Db 74580 CCAACATGATGACAGATGACCTCAATGCTAATCAATTAAGTGAATTAAGT 74521  
Qy 1737 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1796  
Db 74520 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 74461  
Qy 1797 AGACGCTGCAATATGATATTTCTTTATTAACAGAAAGTTCTATTTGCAAAATCTT 1856  
Db 74460 AGACGCTGCAATATGATATTTCTTTATTAACAGAAAGTTCTATTTGCAAAATCTT 74401  
Qy 1857 CTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1916  
Db 74400 CTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 74341  
Qy 1917 GGGCTGAGGAGATGATGATGCAAGTGTATGAGAAAGTTTGGGGTATTAAGCTGT 1976  
Db 74340 GGGCTGAGGAGATGATGATGCAAGTGTATGAGAAAGTTTGGGGTATTAAGCTGT 74281  
Qy 1977 TGCATCTTGAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2036  
Db 74280 TGCATCTTGAATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 74221  
Qy 2037 CACTGGAATTTGGCAATTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2096  
Db 74220 CACTGGAATTTGGCAATTTCAATGATGATGATGATGATGATGATGATGATGATGAT 74161  
Qy 2097 GTCAAGCTTTTGGCAATTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2156  
Db 74160 GTCAAGCTTTTGGCAATTTCAATGATGATGATGATGATGATGATGATGATGATGAT 74101  
Qy 2157 TGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2216  
Db 74100 TGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 74041  
Qy 2217 ACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2276  
Db 74040 ACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73981  
Qy 2277 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2336  
Db 73980 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73921  
Qy 2337 AAAAGTGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2396  
Db 73920 AAAAGTGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73861  
Qy 2397 AAAAGAGAGAAATTTTGGCTTTGAGAGATGAGAGAGATGATGATGATGATGATGAT 2456  
Db 73860 AAAAGAGAGAAATTTTGGCTTTGAGAGATGAGAGAGATGATGATGATGATGATGAT 73801  
Qy 2457 GTTTTAAACATTTCTGCTGAGAGAGAGATGATGATGATGATGATGATGATGATGAT 2516  
Db 73800 GTTTTAAACATTTCTGCTGAGAGAGAGATGATGATGATGATGATGATGATGATGAT 73741

Qy 2517 AATATGCAACCTAGAGAAAGTGCATGAAGGAGAGAGTGTAAATTAATTTTCAATGAT 2576  
Db 73740 AATATGCAACCTAGAGAAAGTGCATGAAGGAGAGAGTGTAAATTAATTTTCAATGAT 73681  
Qy 2577 GTAATGAGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2636  
Db 73680 GTAATGAGAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73621  
Qy 2637 CCTAGAGTCAAGAGAAACAAAGTGTCTCTTCTCTAGTTATGTCAGTGTGCAAGT 2696  
Db 73620 CCTAGAGTCAAGAGAAACAAAGTGTCTCTTCTCTAGTTATGTCAGTGTGCAAGT 73561  
Qy 2697 CCAAAACACCTTCTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCT 2756  
Db 73560 CCAAAACACCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCT 73501  
Qy 2757 CACAAACATCATTTAAACAGGAGGTCATGTCAGAAAGCCATTTCTTTCTCTAGACTT 2816  
Db 73500 CACAAACATCATTTAAACAGGAGGTCATGTCAGAAAGCCATTTCTTTCTCTAGACTT 73441  
Qy 2817 CTATGATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2876  
Db 73440 CTATGATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73381  
Qy 2877 CCACTTGCATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2936  
Db 73380 CCACTTGCATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73321  
Qy 2937 GAATTAAGAGAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2996  
Db 73320 GAATTAAGAGAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73261  
Qy 2997 TATTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3056  
Db 73260 TATTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73201  
Qy 3057 TTTGCTCTTTTCTGAGTGAACCTGCGCATTTATTTCCATTAATGATGAGAGCA 3116  
Db 73200 TTTGCTCTTTTCTGAGTGAACCTGCGCATTTATTTCCATTAATGATGAGAGCA 73141  
Qy 3117 TCTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3176  
Db 73140 TCTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73081  
Qy 3177 AGTAACCTTGAAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3236  
Db 73080 AGTAACCTTGAAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 73021  
Qy 3237 TATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3296  
Db 73020 TATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72961  
Qy 3297 GGGCTATGCTGTAAATCCCAACCTTGGGAAGCCGGGTCGGAGAGATGATGATGATGAT 3356  
Db 72960 GGGCTATGCTGTAAATCCCAACCTTGGGAAGCCGGGTCGGAGAGATGATGATGATGAT 72901  
Qy 3357 GCAATTTGAGCCAGTCAAGGCAACACAGCAAGCCCATATCTAATAAAACAAACAA 3416  
Db 72900 GCAATTTGAGCCAGTCAAGGCAACACAGCAAGCCCATATCTAATAAAACAAACAA 72841  
Qy 3417 ACATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3476  
Db 72840 ACATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 72781  
Qy 3477 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3536  
Db 72780 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72721  
Qy 3537 ACTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3596  
Db 72720 ACTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72661

Oy	3597	TTTTAAAAA	CAAAATTAATTAATTTAAAAACACACACTAGAGATGTTCCAA	3656
Db	72660	TTTTAAAAA	CAAAATTAATTAATTTAAAAACACACACTAGAGATGTTCCAA	72601
Oy	3657	TTGATTAAT	TGGAGTCTAATATCCCTGGAAGTTAAATTTAGAGAGTTCCTC	3716
Db	72600	TTGATTAAT	TGGAGTCTAATATCCCTGGAAGTTAAATTTAGAGAGTTCCTC	72541
Oy	3717	CTCATTTTC	CTAGACGCTGCAATTTGTAAATATCAGAGCTAGAGCAACTAGAGCTCCG	3776
Db	72546	CTCATTTTC	CTAGACGCTGCAATTTGTAAATATCAGAGCTAGAGCAACTAGAGCTCCG	72481
Oy	3777	CACCTCAAA	GTGCTGCTCAAGACACAGACATCAAGTAACTGCGAAAGCTGTAAAT	3836
Db	72486	CACCTCAAA	GTGCTGCTCAAGACACAGACATCAAGTAACTGCGAAAGCTGTAAAT	72421
Oy	3837	GCAGATCTT	AGGCTGCAACCCAGACCTAATCAGAACTCTGCAATTAACAAGTTCT	3896
Db	72420	GCAGATCTT	AGGCTGCAACCCAGACCTAATCAGAACTCTGCAATTAACAAGTTCT	72361
Oy	3897	AGGTGCTCA	CGGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4076
Db	72360	AGGTGCTCA	CGGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	72182
Oy	3957	TTTCATTA	TAAATGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4136
Db	72300	TTTCATTA	TAAATGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	72122
Oy	4017	AGAACTTCA	TAAATGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4196
Db	72241	AGAACTTCA	TAAATGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	72062
Oy	4077	AATGCAAT	TAAATGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4256
Db	72181	AATGCAAT	TAAATGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	72002
Oy	4137	TTTCGAGA	GAAGTCTGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4316
Db	72121	TTTCGAGA	GAAGTCTGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	71942
Oy	4197	AATTCAGA	ATGTAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4376
Db	72061	AATTCAGA	ATGTAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	71882
Oy	4257	AGCTTCTC	CTCCTGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4436
Db	72001	AGCTTCTC	CTCCTGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	71822
Oy	4317	GAATTAAT	TAAATGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4496
Db	71941	GAATTAAT	TAAATGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	71762
Oy	4377	CAACTTGA	CTAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4556
Db	71881	CAACTTGA	CTAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	71702
Oy	4437	GGGAAGAT	GAAGTCTGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4616
Db	71821	GGGAAGAT	GAAGTCTGCAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	71642
Oy	4497	CTGTGAGT	CTAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4676
Db	71761	CTGTGAGT	CTAATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	71602
Oy	4557	CAAAACAG	ATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4736
Db	71701	CAAAACAG	ATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	71582
Oy	4617	ACTTCTGA	TAAATTAATTTGCTAATAGAGAAAGCAACTCTT	4796
Db	71641	ACTTCTGA	TAAATTAATTTGCTAATAGAGAAAGCAACTCTT	71522
Oy	4677	GTAGCTCA	AGTCAATTAATTTGCTAATAGAGAAAGCAACTCTT	4856
Db	4677	GTAGCTCA	AGTCAATTAATTTGCTAATAGAGAAAGCAACTCTT	71402
Oy	4737	GGAGCGA	CTAGATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	4916
Db	71581	GGAGCGA	CTAGATTAATTAATTTGCTAATAGAGAAAGCAACTCTT	71342
Oy	4797	ATTGTGAC	TTGCAATTAATTTGCTAATAGAGAAAGCAACTCTT	4976
Db	71461	ATTGTGAC	TTGCAATTAATTTGCTAATAGAGAAAGCAACTCTT	71282
Oy	4857	AAATTAAT	TAAATTAATTTGCTAATAGAGAAAGCAACTCTT	5036
Db	71401	AAATTAAT	TAAATTAATTTGCTAATAGAGAAAGCAACTCTT	71222
Oy	4917	ATTAAGTT	ATTTGCTAATTTGCTAATAGAGAAAGCAACTCTT	5096
Db	71341	ATTAAGTT	ATTTGCTAATTTGCTAATAGAGAAAGCAACTCTT	71162
Oy	4977	ATCTCTTC	CTTGAATTAATTTGCTAATAGAGAAAGCAACTCTT	5156
Db	71281	ATCTCTTC	CTTGAATTAATTTGCTAATAGAGAAAGCAACTCTT	71102
Oy	5037	CCCTGCTT	GTGCAATTAATTTGCTAATAGAGAAAGCAACTCTT	5216
Db	71221	CCCTGCTT	GTGCAATTAATTTGCTAATAGAGAAAGCAACTCTT	71042
Oy	5097	AGGCGGGA	AGGGAATTAATTTGCTAATAGAGAAAGCAACTCTT	5276
Db	71161	AGGCGGGA	AGGGAATTAATTTGCTAATAGAGAAAGCAACTCTT	70982
Oy	5157	CTTCAAAA	CAATTTGCTAATTTGCTAATAGAGAAAGCAACTCTT	5336
Db	71101	CTTCAAAA	CAATTTGCTAATTTGCTAATAGAGAAAGCAACTCTT	70922
Oy	5217	CAGATACA	GAAGTCTGCAATTAATTTGCTAATAGAGAAAGCAACTCTT	5396
Db	71041	CAGATACA	GAAGTCTGCAATTAATTTGCTAATAGAGAAAGCAACTCTT	70862
Oy	5277	TACAAAAT	CAATTTGCTAATTTGCTAATAGAGAAAGCAACTCTT	5456
Db	70981	TACAAAAT	CAATTTGCTAATTTGCTAATAGAGAAAGCAACTCTT	70802
Oy	5337	GAAGATCA	GAAGTCTGCAATTTGCTAATAGAGAAAGCAACTCTT	5516
Db	70921	GAAGATCA	GAAGTCTGCAATTTGCTAATAGAGAAAGCAACTCTT	70742
Oy	5397	ACTATAGC	TAGCAATTTGCTAATAGAGAAAGCAACTCTT	5576
Db	70861	ACTATAGC	TAGCAATTTGCTAATAGAGAAAGCAACTCTT	70682
Oy	5457	CCGAGCTG	TAGCAATTTGCTAATAGAGAAAGCAACTCTT	5636
Db	70801	CCGAGCTG	TAGCAATTTGCTAATAGAGAAAGCAACTCTT	70622
Oy	5517	AAATTAAT	TAAATTAATTTGCTAATAGAGAAAGCAACTCTT	5696
Db	70741	AAATTAAT	TAAATTAATTTGCTAATAGAGAAAGCAACTCTT	70562
Oy	5577	TCATAGCT	AGTCAATTTGCTAATAGAGAAAGCAACTCTT	5756
Db	70681	TCATAGCT	AGTCAATTTGCTAATAGAGAAAGCAACTCTT	70502
Oy	5637	ACAGAGAT	TAATTTGCTAATAGAGAAAGCAACTCTT	5816
Db	70621	ACAGAGAT	TAATTTGCTAATAGAGAAAGCAACTCTT	70542
Oy	5697	TAAGGAAT	TAATTTGCTAATAGAGAAAGCAACTCTT	5876
Db	70561	TAAGGAAT	TAATTTGCTAATAGAGAAAGCAACTCTT	70512
Oy	5757	ACATCCAG	CGGTAATTTGCTAATAGAGAAAGCAACTCTT	5936
Db	5757	ACATCCAG	CGGTAATTTGCTAATAGAGAAAGCAACTCTT	70452









Db	4677	CCGATATATACAGTATTGGGATATATGAACAATTTATATGGTTCTTGGAAAGTAGG	4730
Oy	4682	TCMACTGAAACCAAAACCAAAACCGCAAAAACCTGTAAACCTTAAGATATAGTCGAGC	4741
Db	4731	TCMACTGAAACCAAAACCAAAACCGCAAAAACCTGTAAACCTTAAGATATAGTCGAGC	4790
Oy	4742	CCACTGAGAGATATAAATAAATAGAAATATTTTATTAACAGGCATTTGAAATATTTG	4801
Db	4791	CGACTGAGAGATATAAATAAATAGAAATATTTTATTAACAGGCATTTGAAATATTTG	4850
Oy	4802	TGCATTGCGAAATTTCTGCAATTAATATTTTCCAAATTTTAATCTTTAAGAAAT	4861
Db	4851	TGCATTGCGAAATTTCTGCAATTAATATTTTCCAAATTTTAATCTTTAAGAAAT	4910
Oy	4862	TACTATATTAATATGTAAATGATCATGTCATGTCTTGAAGTNGAATTTAATCAATATA	4921
Db	4911	TACTATATTAATATGTAAATGATCATGTCATGTCTTGAAGTNGAATTTAATCAATATA	4970
Oy	4922	GGTATATTTCTTTATATTCGGGTCAAGCAAGCTCTTAAGGGGATGAAAGGATATTC	4981
Db	4971	GGTATATTTCTTTATATTCGGGTCAAGCAAGCTCTTAAGGGGATGAAAGGATATTC	5030
Oy	4982	TTTCTTTAGCTGAGAGAAAGATGAGTTCTTAAGTAAATATTAATCAAGCAATTTCCCG	5041
Db	5031	TTTCTTTAGCTGAGAGAAAGATGAGTTCTTAAGTAAATATTAATCAAGCAATTTCCCG	5090
Oy	5042	TCCTTGTATTTAGATTTGTACCAACAACGCGGTGTGGCTGAAGGCAATCGAAGGCG	5101
Db	5091	TCCTTGTATTTAGATTTGTACCAACAACGCGGTGTGGCTGAAGGCAATCGAAGGCG	5150
Oy	5102	GGGAGGGGAGGAAATGATGAAATTAATCAAAACCAAACTTCCCTAAGCGCTCTA	5161
Db	5151	GGGAGGGGAGGAAATGATGAAATTAATCAAAACCAAACTTCCCTAAGCGCTCTA	5210
Oy	5162	CAAAACATTTTAGCCCCGAAATAGTCAAGAAATCTCAAAATCAACCAATTCAGAT	5221
Db	5211	CAAAACATTTTAGCCCCGAAATAGTCAAGAAATCTCAAAATCAACCAATTCAGAT	5270
Oy	5222	ACAAGGAATGTTATGTAGCTGAGAGAGGGGTGAGACTATCAAGCTCAGTTCAAGTAA	5281
Db	5271	ACAAGGAATGTTATGTAGCTGAGAGAGGGGTGAGACTATCAAGCTCAGTTCAAGTAA	5330
Oy	5282	AATCCAGGCTCCGCAATTTAAATCTCGAATGCATTCAGCAGATCCAAATCAACAGG	5341
Db	5331	AATCCAGGCTCCGCAATTTAAATCTCGAATGCATTCAGCAGATCCAAATCAACAGG	5390
Oy	5342	AGATCGAAGTTCAAGATGCGCTCCAGCTCCAAATTTGCCAACACAGATGTGCTACTAT	5401
Db	5391	AGATCGAAGTTCAAGATGCGCTCCAGCTCCAAATTTGCCAACACAGATGTGCTACTAT	5450
Oy	5402	ACGTCGAAGACTCTGAACCGCTGAGAGAGGGGAGAAACAACAGTAGAGAGATGCCCCAG	5461
Db	5451	ACGTCGAAGACTCTGAACCGCTGAGAGAGGGGAGAAACAACAGTAGAGAGATGCCCCAG	5510
Oy	5462	CTGTGTAAGATTCAGATGTTTATGAAGTTTAAAGTATTAAGTATTAAGTATTAAGTAAAT	5521
Db	5511	CTGTGTAAGATTCAGATGTTTATGAAGTTTAAAGTATTAAGTATTAAGTATTAAGTAAAT	5570
Oy	5522	CAAGAAACCGCTCCGCTCTTTGCAAAATATGTATGAAGAGAGAAAGTCCCTAACTTAT	5581
Db	5571	CAAGAAACCGCTCCGCTCTTTGCAAAATATGTATGAAGAGAGAAAGTCCCTAACTTAT	5630
Oy	5582	GTCTGATAGATTGAGCCATATGCTTTTAAAGCCCGCCGCTTTATATCTAATATTAACA	5641
Db	5631	GTCTGATAGATTGAGCCATATGCTTTTAAAGCCCGCCGCTTTATATCTAATATTAACA	5690
Oy	5642	GGATTTTGGATATTTTAAAT	5701
Db	5691	GGATTTTGGATATTTTAAAT	5750
Oy	5702	GAATATGATTTGAACCTTCAACCGGTGCAACCCCTACAGTACTG	5745
Db	5751	TAGAGGCGCTCGACATCTTAAGTAGCTTGCAATCTCGGATCTG	5794



[illegible]

Db	2641	GGAGTCAGAGAGACAAAGATGCCCTTCTCTCAAGTTATGTCAGAGGCCAAAGTCCAA	2760
Qy	2701	AAACACCTTCCTCTTAAAGTACTTTCTTCCCTCCATACAAATCTAAAGCTTCACA	2760
Db	2701	AACACCTTCCTCTTAAAGTACTTTCTTCCCTCCATACAAATCTAAAGCTTCACA	2760
Qy	2761	AACATCATTTAAACAGGCAAGCTCATAGTCAGAAAGGCAATGCTTTCTAGACTCTAT	2820
Db	2761	AACATCATTTAAACAGGCAAGCTCATAGTCAGAAAGGCAATGCTTTCTAGACTCTAT	2820
Qy	2821	GTACGCTTTTATATTAATAATTTTCGCTTAAAGACCTCAAGCTTTGGAAGATTTCCAC	2880
Db	2821	GTACGCTTTTATATTAATAATTTTCGCTTAAAGACCTCAAGCTTTGGAAGATTTCCAC	2880
Qy	2881	CTTGCAATCAAGATATTAATATCATCATCTTGTATAGTACCTTAGTCCCTTAAGAAAT	2940
Db	2881	CTTGCAATCAAGATATTAATATCATCATCTTGTATAGTACCTTAGTCCCTTAAGAAAT	2940
Qy	2941	AAGGATGAACATATAATATTAAGAAATTAATAGTAAATTAATATATGATTTGCCACTTAT	3000
Db	2941	AAGGATGAACATATAATATTAAGAAATTAATAGTAAATTAATATATGATTTGCCACTTAT	3000
Qy	3001	TTTCAGTTGATGCGTATAGCTTGCAAGCTACGAGTTCGTGTAAATTCATAGAGATTTTG	3060
Db	3001	TTTCAGTTGATGCGTATAGCTTGCAAGCTACGAGTTCGTGTAAATTCATAGAGATTTTG	3060
Qy	3061	CCCTCTTTTCTCGGTCMACTGCGCATTTATTTCCATATGCATAGAGCCATCTT	3120
Db	3061	CCCTCTTTTCTCGGTCMACTGCGCATTTATTTCCATATGCATAGAGCCATCTT	3120
Qy	3121	TTTTCATATTAATCTTATTTTAAATTTTGTGCGATTTATTTCTGTCTCTTAGCTTAGTA	3180
Db	3121	TTTTCATATTAATCTTATTTTAAATTTTGTGCGATTTATTTCTGTCTCTTAGCTTAGTA	3180
Qy	3181	ACTTTAGATTTTAAATATACACTATTTGAAATCATAGATACCTTTAAATGAAATATAT	3240
Db	3181	ACTTTAGATTTTAAATATACACTATTTGAAATCATAGATACCTTTAAATGAAATATAT	3240
Qy	3241	TAAATAGCTAGGCTATTAACCTTTTAAATTTTAAATTTTAAATATGATGATGTCGCT	3300
Db	3241	TAAATAGCTAGGCTATTAACCTTTTAAATTTTAAATTTTAAATATGATGATGTCGCT	3300
Qy	3301	CATGCGGTAAATCCCAACATTTTGGGAAGCCGGGTGGGAGATAGCTTAGTCACAG	3360
Db	3301	CATGCGGTAAATCCCAACATTTTGGGAAGCCGGGTGGGAGATAGCTTAGTCACAG	3360
Qy	3361	TTTGAGACGAGTCAAGGCGCAACACGACGAGGCCCATTTCTTAAAAAACAATAACAA	3420
Db	3361	TTTGAGACGAGTCAAGGCGCGCAACACGACGAGGCCCATTTCTTAAAAAACAATAACAA	3420
Qy	3421	AATTAAGCTGGATATGATGTGCTCACTGTAGTCAAGCTACACAGGAAGCTGAGGCGA	3480
Db	3421	AATTAAGCTGGATATGATGTGCTCACTGTAGTCAAGCTACACAGGAAGCTGAGGCGA	3480
Qy	3481	AGATATCACTTGAACCCAGAGAGGTTAAGCTCACTGATATCAATGAAAGGCGCTCATACATC	3540
Db	3481	AGATATCACTTGAACCCAGAGAGGTTAAGCTCACTGATATCAATGAAAGGCGCTCATACATC	3540
Qy	3541	AGTCGGGTGACAGTGCAGAAAGCTGTCTCAAAAAATATTAATTAATTAATCTTT	3600
Db	3541	AGTCGGGTGACAGTGCAGAAAGCTGTCTCAAAAAATATTAATTAATTAATCTTT	3600
Qy	3601	AAAAAACAATAATTAATTAATTTTAAAAACAACAACACTAGAGATGTTGCAATATGA	3660
Db	3601	AAAAAACAATAATTAATTAATTTTAAAAACAACAACACTAGAGATGTTGCAATATGA	3660
Qy	3661	TTATTTGGAGTCTATATCCCTGGAAGTTATTTTAAATTTTAAAGAGATTTCTTCTCA	3720
Db	3661	TTATTTGGAGTCTATATCCCTGGAAGTTATTTTAAATTTTAAAGAGATTTCTTCTCA	3720
Qy	3721	TTTCTAGAGACGTCAATTTGTAATATCAAGCTAAAGGAACATAGGCTGCGCACT	3780
Db	3721	TTTCTAGAGACGTCAATTTGTAATATCAAGCTAAAGGAACATAGGCTGCGCACT	3780

Db	372	TTTCCTAGAGACGTCGAAATTGTAATATTCACAGCTAGAGAGAACCTAGGGCTCGCCACT	3780
Qy	3781	CCAAAGTGTGCTCCAGACACACAGCATCAGATGACCTGCGAAACGTGTTAGAAATGCG	3840
Db	3781	CCAAAGTGTGCTCCAGAGCACAGACGACATCAAGTAACCTGGAAACGTGTAGAAATGCG	3840
Qy	3841	AGCTTAGGCTCAGCCCGACGACCTTACGTAAACAGAAATCTGTAAACAAGATTTCTAGGT	3900
Db	3841	AGCTTAGGCTCAGCCCGACGACCTTACGTAAACAGAAATCTGTAAACAAGATTTCTAGGT	3900
Qy	3901	GGCTCAGCGGCATATTAACCTTGAGAGCTCTGACCTAGAAATCTTCACTCCACCTTC	3960
Db	3901	GGCTCAGCGGCATATTAACCTTGAGAGCTCTGACCTAGAAATCTTCACTCCACCTTC	3960
Qy	3961	ATTATTAATGGAATCACTTGGCTGTGTGTACAGAGAAATGATTATTTTAATTTTCAGAA	4020
Db	3961	ATTATTAATGGAATCACTTGGCTGTGTGTACAGAGAAATGATTATTTTAATTTTCAGAA	4020
Qy	4021	CCCTCTAATTAGCTCATATTAATTTGCTAAATAGAGGAGAAAGCCAACTCTTAACT	4080
Db	4021	CCCTCTAATTAGCTCATATTAATTTGCTAAATAGAGGAGAAAGCCAACTCTTAACT	4080
Qy	4081	GCAATTAACGAATCTATAATTAATAGTAAAGCAATCCCTTTAAGTTTACATTTTG	4140
Db	4081	GCAATTAACGAATCTATAATTAATAGTAAAGCAATCCCTTTAAGTTTACATTTTG	4140
Qy	4141	TGAGAGCACTGCTTTGATTTGGCTGGGCTCAGGCTGGCTGTTGTGTGAATTTTCACATT	4200
Db	4141	TGAGAGCACTGCTTTGATTTGGCTGGGCTCAGGCTGGCTGTTGTGTGAATTTTCACATT	4200
Qy	4201	CACAGATGTTAAGCCGCTCTCGGCTAGTAAAGGAGAGAAATTCAGTTTAATTACT	4260
Db	4201	CACAGATGTTAAGCCGCTCTCGGCTAGTAAAGGAGAGAAATTCAGTTTAATTACT	4260
Qy	4261	TCTCCCTCATCTCGGCTGAAGCAACAAATTAATTTTATGAAACACATTTTGACT	4320
Db	4261	TCTCCCTCATCTCGGCTGAAGCAACAAATTAATTTTATGAAACACATTTTGACT	4320
Qy	4321	TAAATTTACTTACAGGGAATGTCAAAATTTCTCTGAAGGCTTAAAGATGTCTCAACAC	4380
Db	4321	TAAATTTACTTACAGGGAATGTCAAAATTTCTCTGAAGGCTTAAAGATGTCTCAACAC	4380
Qy	4381	TTTGACATCTACGTAGTGTACCTAATTTTACAGGTGTCTCGTGACTAGAGGCGTGAAGGA	4440
Db	4381	TTTGACATCTACGTAGTGTACCTAATTTTACAGGTGTCTCGTGACTAGAGGCGTGAAGGA	4440
Qy	4441	AGATGTGAACCTCAGCATGTTAGTACGCTAGTAACAAGACGCTTTTATTTTCCCTG	4500
Db	4441	AGATGTGAACCTCAGCATGTTAGTACGCTAGTAACAAGACGCTTTTATTTTCCCTG	4500
Qy	4501	TTGGAATCTAATCTTACGTACCTCTGTAAATCTAATTTCAATTTCCAAATCCACAA	4560
Db	4501	TTGGAATCTAATCTTACGTACCTCTGTAAATCTAATTTCAATTTCCAAATCCACAA	4560
Qy	4561	ACCAGATTAAGTTTACAGCCCATATTTACAGAAAGGAATTAATTTTGTGTGTAGACT	4620
Db	4561	ACCAGATTAAGTTTACAGCCCATATTTACAGAAAGGAATTAATTTTGTGTGTAGACT	4620
Qy	4621	TCTGTATTTACAGCTATTTGGGAATTAAGAACTTTTAATGTTTCTTTGCAATAG	4680
Db	4621	TCTGTATTTACAGCTATTTGGGAATTAAGAACTTTTAATGTTTCTTTGCAATAG	4680
Qy	4681	GTCAGATCAAGCAAAACCAAAACAGCAAAACGTGAAGACATTAAGATAGTGAAG	4740
Db	4681	GTCAGATCAAGCAAAACCAAAACAGCAAAACGTGAAGACATTAAGATAGTGAAG	4740
Qy	4741	CCGACTGAGAGATTAATTAATTAATTAATTAATTTTAAACGGCAATTTGAAATTAATTT	4800
Db	4741	CCGACTGAGAGATTAATTAATTAATTAATTAATTTTAAACGGCAATTTGAAATTAATTT	4800
Qy	4801	GTCGACTCAGAAATTTCTACAAATTAATTAATTTCCAAATTTTAATCTTTTAAGAAA	4860
Db	4801	GTCGACTCAGAAATTTCTACAAATTAATTAATTTCCAAATTTTAATCTTTTAAGAAA	4860

QY	4881	TTACTATATATATATGAAAGTAAATCTGCATATGCTTTACGCTAGCAATATTTAACTCATPA	4920
DB	4881	TTACTATATATATATGTAATCATCTGCACTGCTTTAGAGTACAAATATTTAACTCATPA	4920
QY	4921	AGGTTATATTTCTTTATATCGGCTCAGCGAAAGCTTCTTAAGGAGATGCAAGGGAATATCT	4980
DB	4921	AGGTTATATTTCTTTATATCGGCTCAGCGAAAGCTTCTTAAGGAGATGCAAGGGAATATCT	4980
QY	4981	CTTCTCTTAGCTGAGAGAGAGAGCTCTAAGTTAAATATATATCAAGAAATTTCCCT	5040
DB	4981	CTTCTCTTAGCTGAGAGAGAGAGCTCTAAGTTAAATATATATCAAGAAATTTCCCT	5040
QY	5041	GTCCTTGCTATTTGAGATTTGACCAACAACGCGGTTGCTGAAAGGAAACTGAAGG	5100
DB	5041	GTCCTTGCTATTTGAGATTTGACCAACAACGCGGTTGCTGAAAGGAAACTGAAGG	5100
QY	5101	CGGGAGGAGAGGAAATATGATGAAAAAACAACAAACAACTTCCTTAAGCAGCTCT	5160
DB	5101	CGGGAGGAGAGGAAATATGATGAAAAAACAACAAACAACTTCCTTAAGCAGCTCT	5160
QY	5161	ACAAAACATTTTAGCCCCAGAAATATGTCACAGAAATCTCTCAATCAAAACGATATCCGA	5220
DB	5161	ACAAAACATTTTAGCCCCAGAAATATGTCACAGAAATCTCTCAATCAAAACGATATCCGA	5220
QY	5221	TACAAGAAAGTATATATGATGCTGAGAGAGGCTGACATCATAGCTCACTCAGTTCAATACA	5280
DB	5221	TACAAGAAAGTATATATGATGCTGAGAGAGGCTGACATCATAGCTCACTCAGTTCAATACA	5280
QY	5281	AAAGTCCAGGCTGCTGAATATTAACCTGTATGCAATTCATGCGACATTCATCAATCAGACA	5340
DB	5281	AAAGTCCAGGCTGCTGAATATTAACCTGTATGCAATTCATGCGACATTCATCAATCAGACA	5340
QY	5341	GAGATCAGAAAGTTCAAGATGCTCTCAGAGCTCCAAATTTGCAACAACAGATGTGCTACTA	5400
DB	5341	GAGATCAGAAAGTTCAAGATGCTCTCAGAGCTCCAAATTTGCAACAACAGATGTGCTACTA	5400
QY	5401	TAGCTCAAGAGACTCTGAAGCCGTGAGAGAGGAGGAAACAACATGAGAGAGATGCCA	5460
DB	5401	TAGCTCAAGAGACTCTGAAGCCGTGAGAGAGGAGGAAACAACATGAGAGAGATGCCA	5460
QY	5461	GCTGCTAGCAATCGAGTGTATTATCAAGTTTACTCAATTTGATGAATCTCATTTGCTAAA	5520
DB	5461	GCTGCTAGCAATCGAGTGTATTATCAAGTTTACTCAATTTGATGAATCTCATTTGCTAAA	5520
QY	5521	TCAGAAAGCGCTCCGCTCTTTGCAAAATATGATGAAGAGAGAGAGTCCCTAAACTCTA	5580
DB	5521	TCAGAAAGAGCGCTCCGCTCTTTGCAAAATATGATGAAGAGAGAGTCCCTAAACTCTA	5580
QY	5581	TGTCGTATAGCATTTGAAACCATTTGCTTTTAGGCTCCCGGCTTTATATCTATATATAC	5640
DB	5581	TGTCGTATAGCATTTGAAACCATTTGCTTTTAGGCTCCCGGCTTTATATCTATATATAC	5640
QY	5641	AGGATATTTGTATATTTATATATATATTTGTTCTCCCT 5676	
DB	5641	AGGATATTTGTATATTTATATATATATTTGTTCTCCCT 5676	

RESULT 5					
BD093091/c	BD093091	5676 bp	DNA	linear	PAT 27-AUG-2002
LOCUS					
DEFINITION	Gene encoding promoter region of tumor suppressor gene p51.				
ACCESSION	BD093091				
VERSION	BD093091.1	GI:22638679			
KEYWORDS	WO 0100818-A/3.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 (bases 1 to 5676)				
AUTHORS	Sakai,T., Kagawa,S., Sato,T., Suenaga,Y. and Fujii,H.				
TITLE	Gene encoding promoter region of tumor suppressor gene p51				
JOURNAL	Patent: WO 0100818-A 3 04-JAN-2001.				

**BEST AVAILABLE COPY**





Db 3876 GGTGAAATATAGAAATTTCTCTTAAACAGAAAGTTCTATGAAAAACAAAGTTCTTAG 3817  
QY 1861 AAAATATACATTAATCTATTAATAAACAAGCAAGTCAAGATTTGTTAGGCCAGGAG 1920  
Db 3816 AAAATATACATTAATCTATTAATAAACAAGCAAGTCAAGATTTGTTAGGCCAGGAG 3757  
QY 1921 TAGAGGAGATCGATTGCAAGTGTATAGAGAAAGTTTGGGGTAATAGGGTTGTTGCA 1980  
Db 3756 TAGAGGAGATCGATTGCAAGTGTATAGAGAAAGTTTGGGGTAATAGGGTTGTTGCA 3697  
QY 1981 ATCTGATTTGCAAGAGGCTACGAGTGTCTAATAGTCACTCCCTCAAGCAAGTCACT 2040  
Db 3696 ATCTGATTTGCAAGAGGCTACGAGTGTCTAATAGTCACTCCCTCAAGCAAGTCACT 3637  
QY 2041 TGCAATTTGCGAATTTCAATTGTAATTTATACCTCAATAAGTAACTCTAAGAGTCA 2100  
Db 3636 TGCAATTTGCGAATTTCAATTGTAATTTATACCTCAATAAGTAACTCTAAGAGTCA 3577  
QY 2101 AGTGTGTTGTGAAATTTATTTTAAATCAAGTTCAGTCAATTAATAGATGATTTTGA 2160  
Db 3576 AGTGTGTTGTGAAATTTATTTTAAATCAAGTTCAGTCAATTAATAGATGATTTTGA 3517  
QY 2161 AATACATTAACATGTTATTCATTCATAGTCAATTTTGTCTAGCTCTCAAAAC 2220  
Db 3516 AATACATTAACATGTTATTCATTCATAGTCAATTTTGTCTAGCTCTCAAAAC 3457  
QY 2221 AGAGATGAATTAAGATAGCAAGCTCCCTCAAGCTGTTCAATCCAGTACAGAGATG 2280  
Db 3456 AGAGATGAATTAAGATAGCAAGCTCCCTCAAGCTGTTCAATCCAGTACAGAGATG 3397  
QY 2281 AGTCTATTTCAAAATAGTACACTCCAGAGAAAGTTATAGGAGCTTTACCAAAA 2340  
Db 3396 AGTCTATTTCAAAATAGTACACTCCAGAGAAAGTTATAGGAGCTTTACCAAAA 3337  
QY 2341 GTGAGATTAATTAATTAATAGTACAGTGAAGTGGGAAAGTTTCTTTATAGTGA 2400  
Db 3336 GTGAGATTAATTAATTAATAGTGAAGTGGGAAAGTTTCTTTATAGTGA 3277  
QY 2401 GAGGAGAAATTTGTTGCTTTAGAGATAGCAAGATGTAATAGCAATGCAATTT 2460  
Db 3276 GAGGAGAAATTTGTTGCTTTAGAGATAGCAAGATGTAATAGCAATGCAATTT 3217  
QY 2461 TAAACATCTCTGATGAGGAGCAAGTATGATCCAGAGCAAGAGCAAGCAAGAA 2520  
Db 3216 TAAACATCTCTGATGAGGAGCAAGTATGATCCAGAGCAAGAGCAAGAA 3157  
QY 2521 TGCAACTAGAGAAAGTCAAGAGGAGCAAGTGTAAATTAATTTTCAATGAATGA 2580  
Db 3156 TGCAACTAGAGAAAGTCAAGAGGAGCAAGTGTAAATTAATTTTCAATGAATGA 3097  
QY 2581 GTGAGAAAGATTTGTATCATAGACACTGAGTTTGGAGAGTCAATGTTCTAGCTCA 2640  
Db 3096 GTGAGAAAGATTTGTATCATAGACACTGAGTTTGGAGAGTCAATGTTCTAGCTCA 3037  
QY 2641 GGAATCAAGAGAAACAAAGTGTCCCTTCTCTAGTAAAGTCAAGTGTCAAGTCA 2700  
Db 3036 GGAATCAAGAGAAACAAAGTGTCCCTTCTCTAGTAAAGTGTCAAGTGTCA 2977  
QY 2701 AAGCTCTTCTCTCTTAAAGTACTTTCTTCTCTCTCAATCAATCTAAGTCTTCA 2760  
Db 2976 AAGCTCTTCTCTCTTAAAGTACTTTCTTCTCTCTCAATCAATCTAAGTCTTCA 2917  
QY 2761 AACATCTTTAAACAGGAGGTCATGTGCAAGAAAGCAATGCTTTTCTAGACTTCA 2820  
Db 2916 AACATCTTTAAACAGGAGGTCATGTGCAAGAAAGCAATGCTTTTCTAGACTTCA 2857  
QY 2821 GAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2880  
Db 2856 GAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2797  
QY 2881 CTTCGACATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940

Db 2796 CTTCGACATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2737  
QY 2941 AAGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3000  
Db 2736 AAGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2677  
QY 3001 TTTCATCTGATCGGTATAGGTTGCAATGCTACGTGTTCTTGAATTTCAAGAGTTG 3060  
Db 2676 TTTCATCTGATCGGTATAGGTTGCAATGCTACGTGTTCTTGAATTTCAAGAGTTG 2617  
QY 3061 CCTCTTTTCTCGGTCACTTCGCAATTAATTTTCCATATAGCAATAGAGCAATCTT 3120  
Db 2616 CCTCTTTTCTCGGTCACTTCGCAATTAATTTTCCATATAGCAATAGAGCAATCTT 2857  
QY 3121 TTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180  
Db 2556 TTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2497  
QY 3181 ACTTAGAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3240  
Db 2496 ACTTAGAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2437  
QY 3241 TAAATAGGTTAGGCTATTAACCTTTTAAATTTTAAATTAATTAATTAATTAATTA 3300  
Db 2436 TAAATAGGTTAGGCTATTAACCTTTTAAATTTTAAATTAATTAATTAATTAATTA 2377  
QY 3301 CATGCTGTAAATTCACACCTTTGGGAAAGCGGGTCGGAGATAGCTGATTCACAG 3360  
Db 2376 CATGCTGTAAATTCACACCTTTGGGAAAGCGGGTCGGAGATAGCTGATTCACAG 2317  
QY 3361 TTTCAGACCAATCGGCAACAGCAAGCAAGCCATCTAATAAACAACAACAACA 3420  
Db 2316 TTTCAGACCAATCGGCAACAGCAAGCAAGCCATCTAATAAACAACAACAACA 2257  
QY 3421 AATTAAGGTTAGGTTGTTGCTCACTGTAGTCCAACTACAGAGAGTGGGCA 3480  
Db 2256 AATTAAGGTTAGGTTGTTGCTCACTGTAGTCCAACTACAGAGAGTGGGCA 2197  
QY 3481 AGATCACTTGAAGCCAGAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAG 3540  
Db 2196 AGATCACTTGAAGCCAGAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAG 2137  
QY 3541 AGTCTGGGTCAAGTGCAGAGAGCTGTCTCAAAATTAATTAATTAATTAATTAAT 3600  
Db 2136 AGTCTGGGTCAAGTGCAGAGAGCTGTCTCAAAATTAATTAATTAATTAATTAAT 2077  
QY 3601 AAAAAACAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3660  
Db 2076 AAAAAACAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2017  
QY 3661 TTAATTTGGAGGTCTATATCCCTGGAAAGTTAATTAATTAATTAATTAATTAATTA 3720  
Db 2016 TTAATTTGGAGGTCTATATCCCTGGAAAGTTAATTAATTAATTAATTAATTAAT 1957  
QY 3721 TTTCTGAGAGCTCGAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3780  
Db 1956 TTTCTGAGAGCTCGAATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1897  
QY 3781 CCAAGTGTGTGTCAGAGCAAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 3840  
Db 1896 CCAAGTGTGTGTCAGAGCAAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1837  
QY 3841 AGCTTAGAGGCTGACCCGAGACTCTGAAACAGAACTGCAATTAACAGATTTAGGT 3900  
Db 1836 AGCTTAGAGGCTGACCCGAGACTCTGAAACAGAACTGCAATTAACAGATTTAGGT 1777  
QY 3901 GCCTCAGGGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3960  
Db 1776 GCCTCAGGGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1717  
QY 3961 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020  
Db 1716 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1657

BEST AVAILABLE COPY



**BEST AVAILABLE COPY**

QY	4021	CCCTCTATTATAGCATCTATATTTGCTATATAGAGGGAGAAAAGCCAAATCTTTAACT	4080
Db	1656	CCCTCTATTATAGCATCTATATTTGCTATATAGAGGGAGAAAAGCCAAATCTTTAACT	1597
QY	4081	GCAATTTAACAATCTATATATTAATTAAGTAAAGCAATCTTCCCTTTAGTTTACATTTTG	4140
Db	1596	GCAATTTAACAATCTATATATTAATTAAGTAAAGCAATCTTCCCTTTAGTTTACATTTTG	1537
QY	4141	TGAGAGAGCTGTTGATTTGGCTGGGGGCTCAGGCCGGCTGTGTTGTGTAATTTCACAAT	4200
Db	1536	TGAGAGAGCTGTTGATTTGGCTGGGGGCTCAGGCCGGCTGTGTTGTGTAATTTCACAAT	1477
QY	4201	CACAGATGTATACCCCTCTGGGCTATGTAAAGAAAGAAATGTCAAGTTTTAAATAGCT	4260
Db	1476	CACAGATGTATACCCCTCTGGGCTATGTAAAGAAAGAAATGTCAAGTTTTAAATAGCT	1417
QY	4261	TCCTCCTTCCATCTCGGCTGAAGCAACAATAAATTTTATATGAAACAATTGAGT	4320
Db	1416	TCCTCCTTCCATCTCGGCTGAAGCAACAATAAATTTTATATGAAACAATTGAGT	1357
QY	4321	TGATTTTACCTAACAGGAAATGTCAAAATTTCTCTGAAGAGGCTTTAGATGTCTCAAC	4380
Db	1356	TGATTTTACCTAACAGGAAATGTCAAAATTTCTCTGAAGAGGCTTTAGATGTCTCAAC	1297
QY	4381	TTTGACATCTACGTATGTCACTATTTTAAAGGTGTCTGTAGCTAGGGGGTGAAGGA	4440
Db	1296	TTTGACATCTACGTATGTCACTATTTTAAAGGTGTCTGTAGCTAGGGGGTGAAGGA	1237
QY	4441	AGATGTGAATCAGCAAGATTTAGTACGGTATAGATACAGAGAGGTTTATTTTCCCCCTG	4500
Db	1236	AGATGTGAATCAGCAAGATTTAGTACGGTATAGATACAGAGAGGTTTATTTTCCCCCTG	1177
QY	4501	TTGGAGCTATCTTACTAGAGCTTCTGATATCATTTTCAATTTTCCAAATCCAGAA	4560
Db	1176	TTGGAGCTATCTTACTAGAGCTTCTGATATCATTTTCAATTTTCCAAATCCAGAA	1117
QY	4561	ACCAAGATTAAGTTTACAGCCCATTTCTGAAAGGAAATAATTTATTTGTGTAGACTT	4620
Db	1116	ACCAAGATTAAGTTTACAGCCCATTTCTGAAAGGAAATAATTTATTTGTGTAGACTT	1057
QY	4621	TCCTGATATTAACATCTGATTTGGGAATATATGACAAATTTATGTTTCCCTTCGAAAGT	4680
Db	1056	TCCTGATATTAACATCTGATTTGGGAATATATGACAAATTTATGTTTCCCTTCGAAAGT	997
QY	4681	GTCAAGTCAAGCAAAACCAAAAACAGCAAAAACTGTAAACATTAAGATAGTGGAG	4740
Db	996	GTCAAGTCAAGCAAAACCAAAAACAGCAAAAACTGTAAACATTAAGATAGTGGAG	937
QY	4741	CCGACTAGAGATTAATTAATTAATTAATTTTATTAACAGGCAATTTGAAATTAAT	4800
Db	936	CCGACTAGAGATTAATTAATTAATTAATTTTATTAACAGGCAATTTGAAATTAAT	877
QY	4801	GTGCACCTCGAATATTTCAACAAATAATATATTTTCAATTTTAAATCTTTAAGAAA	4860
Db	876	GTGCACCTCGAATATTTCAACAAATAATATATTTTCAATTTTAAATCTTTAAGAAA	817
QY	4861	TTACTATTAATATATGTAATTAATTAATTAATTAATTTTATTAACAGGCAATTTTAACTAATTA	4920
Db	816	TTACTATTAATATATGTAATTAATTAATTAATTTTATTAACAGGCAATTTTAACTAATTA	757
QY	4921	AGGTTATTTTCTTTATTTGGGCTGAGGAAAGCTTCTTAAGGGAATGTGAAGGGAATATCT	4980
Db	756	AGGTTATTTTCTTTATTTGGGCTGAGGAAAGCTTCTTAAGGGAATGTGAAGGGAATATCT	697
QY	4981	CTTTCTCTAGCTGAGCAAGAGTGAAGTCTTAAATTAATTAATTAATTTTCCCT	5040
Db	696	CTTTCTCTAGCTGAGCAAGAGTGAAGTCTTAAATTAATTAATTAATTTTCCCT	637
QY	5041	GTCTTTGCTAATTTGAGATTTGTGACCAACAACAGGCGGTTGGCTGAAAGGGAATCTGAAGGG	5100
Db	636	GTCTTTGCTAATTTGAGATTTGTGACCAACAACAGGCGGTTGGCTGAAAGGGAATCTGAAGGG	577

QY	5101	CGGGAGGGAGGGAAATGATGAAAAA	CAAAACAAACAAACCTCCCTAAGCAGCTC	5166
Db	576	CGGGAGGGAGGAAATGATGAAAAA	CAAAACAAACAAACCTCCCTAAGCAGCTC	517
QY	5161	ACAAACATTTTACCCCGAATAAGT	CAAGAAATCTCAATCAACAGTATCCGA	5220
Db	516	ACAAACATTTTACCCCGAATAAGT	CAAGAAATCTCAATCAACAGTATCCGA	457
QY	5221	TACAAAGAAAGTTATGTAGCTGAG	AGGGTGGACATCATCAGCTCAGTTCAGTTACA	5280
Db	456	TACAAAGAAAGTTATGTAGCTGAG	AGGGTGGACATCATCAGCTCAGTTCAGTTACA	397
QY	5281	AAAGTCAGAGCTGTGAAATTTAA	ACTGTGATCCATTCATGCGAGATTCACAGCA	5340
Db	396	AAAGTCAGAGCTGTGAAATTTAA	ACTGTGATCCATTCATGCGAGATTCACAGCA	337
QY	5341	GAAATCAGAAAGTTCAAGATGCTC	CAAGCTCCAAATTGCAACAAAGTGTGGCTACTA	5400
Db	336	GAAATCAGAAAGTTCAAGATGCTC	CAAGCTCCAAATTGCAACAAAGTGTGGCTACTA	277
QY	5401	TACGTCAAGAACTTGAAACCCCTG	ACAGAGGGGAGAACACACAGTAGAGAGATGCCA	5460
Db	276	TACGTCAAGAACTTGAAACCCCTG	ACAGAGGGGAGAACACACAGTAGAGAGATGCCA	217
QY	5461	GCTGTGAGATCGAGTGTATGAA	AGTTTAACTCATCATGAACTCATGTGGCTAAAA	5520
Db	216	GCTGTGAGATCGAGTGTATGAA	AGTTTAACTCATCATGAACTCATGTGGCTAAAA	157
QY	5521	TCAGAAAGCTCCCGCTCTTTCG	AAATATGTATGAAGAGAGAAATGCTTAACTCTTA	5580
Db	156	TCAGAAAGCTCCCGCTCTTTCG	AAATATGTATGAAGAGAGAAATGCTTAACTCTTA	97
QY	5581	TGCTGTAGCATTTGACCTATATG	CTTTTAGCCTCCCGCTTATATCTATATATACAC	5640
Db	96	TGCTGTAGCATTTGACCTATATG	CTTTTAGCCTCCCGCTTATATCTATATATACAC	37
QY	5641	AGGATTTGTGTATTTTATATAT	ATGTTCTCCGT 5676	
Db	36	AGGATTTGTGTATTTTATATAT	ATGTTCTCCGT 1	

[illegible]

ORIGINAL RESEARCH  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 5676)  
Sakai, T., Kagaya, S., Sato, T., Sukegawa, Y., and Fujii, H.  
Gene encoding promoter region of tumor suppressor gene p51  
Patent: WO 0100818-A 4 04-JAN-2001.

COMMENT OS Homo sapiens (human)

PF	28-JUN-2000	WC	2000JPO04281	
PR	29-JUN-1999	JP	99P	183195
PI	TOSHITSUKU SAKAI,SHIGEHIDE KAKAYA,TAKAMICHI SATO,YOSHITAKU			PI
	SOKENKA,			
PI	HIDEKI FUJII			
PC	CI2N15/10.CI2N5/10.CI201/68.A61K48/00.A61K45/00.A61P43/CO.			PC
	A61P35/00			

```

PH      Key      Location/Qualifiers
FEATURES
source :      1..5676
              /organism="Homo sapiens"

```



3636	TGGAATTGGCCAAATTTCATTGTATGTAAATTAATTAACCTCAATAAAGTAACCTAGAGGTCA	3577	
QY	2101	AGTGTGTTTGGAATTAATTTTAAATCAGTGTGCATTAATTAATGAGATGATTTTGTGA	2160
Db	3576	ACGTGTTTGTGTGAATTAATTTTAAATCAGTGTGCATTAATTAATGAGATGATTTTGTGA	3517
QY	2161	AATACATTAACATGTATTAATCATTCATTAAGTGTGCATATTTTGTGTAGTCTGTGAACAC	2220
Db	3516	AATACATTAACATGTATTAATCATTCATTAAGTGTGCATATTTTGTGTAGTCTGTGAACAC	3457
QY	2221	AAGATGGAATTAAGATTAAGCAAGCTGTGCTCCATAGCTGTTCACATCCAGTACAGAGGATG	2280
Db	3456	AAGATGGAATTAAGATTAAGCAAGCTGTGCTCCATAGCTGTTCACATCCAGTACAGAGGATG	3397
QY	2281	AGTCTATTCAAAATAAGTACACTCCAGAGAAAGTTATATAGTGTACTTAACAAAAA	2340
Db	3396	AGTCTATTCAAAATAAGTACACTCCAGAGAAAGTTATATAGTGTACTTAACAAAAA	3337
QY	2341	GTGCAATATATATATATGTATAGAGACAGTACAGATGCGGAGAGTGTCTTTATGTGAAAAA	2400
Db	3336	GTGCAATATATATATATGTATAGAGACAGTACAGATGCGGAGAGTGTCTTTATGTGAAAAA	3277
QY	2401	GAGGAGGAATTTTGGTCTTTGAAAGATGAGGAAATGTGAAATATGCGCAATAGAGTTT	2460
Db	3276	GAGGAGGAATTTTGGTCTTTGAAAGATGAGGAAATGTGAAATATGCGCAATAGAGTTT	3217
QY	2461	TAAACATNTCTGTGTGAGGGCAGAAATATGATCCAGGACAGAGCAACGAAAAATA	2520
Db	3216	TAAACATNTCTGTGTGAGGGCAGAAATATGATCCAGGACAGAGCAACGAAAAATA	3157
QY	2521	TGCACCTTAGAGGAATAGTGCATGGAAGGGAGAGAGTTGTATTAATTTTCAATGAAATGA	2580
Db	3156	TGCACCTTAGAGGAATAGTGCATGGAAGGGAGAGAGTTGTATTAATTTTCAATGAAATGA	3097
QY	2581	GTGAGGAATTTGTATCATTAAGACACTGAGTTTGGAGAGTGCATGTTCTTGCTCTCA	2640
Db	3096	GTGAGGAATTTGTATCATTAAGACACTGAGTTTGGAGAGTGCATGTTCTTGCTCTCA	3037
QY	2641	GGAGTCAAGAAACAAAGTGTCCCTTCTCTTAACGTTATGCTGAGTGTCCAGTCCAA	2700
Db	3036	GGAGTCAAGAAACAAAGTGTCCCTTCTCTTAACGTTATGCTGAGTGTCCAGTCCAA	2977
QY	2701	AACACCTTCTCCCTTAAGACATTTCTTCTCCCTCCATACAAATCTAAAGTCTTACA	2760
Db	2976	AACACCTTCTCCCTTAAGACATTTCTTCTCCCTCCATACAAATCTAAAGTCTTACA	2917
QY	2761	AACATCATTTAAACAGGCGAGTCAATGTGCAGAAAGCAATGCTTTTCTGACATTCAT	2820
Db	2916	AACATCATTTAAACAGGCGAGTCAATGTGTGCAGAAAGCAATGCTTTTCTGACATTCAT	2857
QY	2821	GTAGCTATTAATTAATCAATTTCTGCTCTAAAGAGACTTAAGTCTTGGAAAGTTTCCAC	2880
Db	2856	GTAGCTATTAATTAATCAATTTCTGCTCTAAAGAGACTTAAGTCTTGGAAAGTTTCCAC	2797
QY	2881	CTTGCACATGAAAGATATTAATCAGTCATTTGTATATGTAACCTTAATCCCTTAAGAGAT	2940
Db	2796	CTTGCACATGAAAGATATTAATCAGTCATTTGTATATGTAACCTTAATCCCTTAAGAGAT	2737
QY	2941	AAGAGTAACATTAATTAATTAAGAGTAATATGTATATTAATTAATATGATGCCATTAAT	3000
Db	2736	AAGAGTAACATTAATTAATTAAGAGTAATATGTATATTAATTAATATGATGCCATTAAT	2677
QY	3001	TTTCACTGTAAATCGTATAGTGTGAGTGCATAGTGTCTGCTGTGAATTTCTAGAGATTGG	3060
Db	2676	TTTCACTGTAAATCGTATAGTGTGAGTGCATAGTGTCTGCTGTGAATTTCTAGAGATTGG	2617
QY	3061	CCCTCTTTTCCGTGGTCACTGTGCGCATTTATTTCCATTAATGCAATAGAGGCAATCTT	2555
Db	2616	CCCTCTTTTCCGTGGTCACTGTGCGCATTTATTTCCATTAATGCAATAGAGGCAATCTT	2497
QY	3121	TTTCAATTAATTAATTAATTAATTTGTGCGCATTAATTTCTGTGCTCTTAGCTTAATGA	3180
Db	2556	TTTCAATTAATTAATTAATTAATTTGTGCGCATTAATTTCTGTGCTCTTAGCTTAATGA	2497

QY	3181	ACTTTAGCAATTTTAAATTAACAACATATGAAATCAATGACATAGCTTTAAATGAAATTAAT	3240
Db	2496	ACTTTAGCAATTTTAAATTAACAACATATGAAATCAATGACATAGCTTTAAATGAAATTAAT	2437
QY	3241	TAAATAGCTTAGGCTATTAACCTTTAAATTTTAAAAAATATGATGATGTGTGGGCT	3300
Db	2436	TAAATAGCTTAGGCTATTAACCTTTAAATTTTAAAAAATATGATGATGTGTGGGCT	2377
QY	3301	CATGCGCTGAATCCCAACACTTTCGGAAAGCCCGGCTCCGAGAGATAGCTTGAGTCAGACAG	3360
Db	2276	CATGCGCTGAATCCCAACACTTTCGGAAAGCCCGGCTCCGAGAGATAGCTTGAGTCAGACAG	2217
QY	3361	TTTGAAACCAAGTCAGGGGCAACAGAGAAAGCCCATATCTAAAAAACAATAACAAAACA	3420
Db	2316	TTTGAAACCAAGTCAGGGGCAACAGAGAAAGCCCATATCTAAAAAACAATAACAAAACA	2257
QY	3421	AATTACCTGGATATGGTTGTGCTCACTGTATGTCCAAAGCTATCAAGCAAGCTGAGGAGA	3480
Db	2256	AATTACCTGGATATGGTTGTGCTCACTGTATGTCCAAAGCTATCAAGCAAGCTGAGGAGA	2197
QY	3481	AGATCACTTGAGCCCGAGAGGTTGAGGCTGCAATGATCCATGAAACGGCTGCTCAACTC	3540
Db	2196	AGATCACTTGAGCCCGAGAGGTTGAGGCTGCAATGATCCATGAAACGGCTGCTCAACTC	2137
QY	3541	AGTCTGGGTGACAGTGCAGAAAGGCTGCTCAAAAATATATTAATTAATTAATTAATCTTT	3600
Db	2136	AGTCTGGGTGACAGTGCAGAAAGGCTGCTCAAAAATATATTAATTAATTAATTAATCTTT	2077
QY	3601	AAAAACAAAAATTAATTAATTTTAAAAACAACAACATAGAGATGTTTGCAAAATGA	3660
Db	2076	AAAAACAAAAATTAATTAATTTTAAAAACAACAACATAGAGATGTTTGCAAAATGA	2017
QY	3661	TTATTTGGAGCTATATCCCTGGAAATTTTAAATTTTAAAGAAAGTCTTCCCA	3720
Db	2016	TTATTTGGAGCTATATCCCTGGAAATTTTAAATTTTAAAGAAAGTCTTCCCA	1957
QY	3721	TTTCTTGAGAGAGCTGCAATTTGTAAATATCAAGCTTAAAGAAACATAGGCTCGCACT	3780
Db	1956	TTTCTTGAGAGAGCTGCAATTTGTAAATATCAAGCTTAAAGAAACATAGGCTCGCACT	1897
QY	3781	CCAAAGTGCTGCCAAGACCAAGCAATCAAGTAACTTGGAACGCTGTGAATAATGAC	3840
Db	1896	CCAAAGTGCTGCCAAGACCAAGCAATCAAGTAACTTGGAACGCTGTGAATAATGAC	1837
QY	3841	AGCTTAAAGGCTCAACCCAGACCTTCAATCAAGCAATCTCATTAACAAGATTTCTAGGT	3900
Db	1836	AGCTTAAAGGCTCAACCCAGACCTTCAATCAAGCAATCTCATTAACAAGATTTCTAGGT	1777
QY	3901	GCTCAAGGGCACTTAAACTTGAAGAGCTCTCACTAATCTTCACTCACTCTC	3960
Db	1776	GCTCAAGGGCACTTAAACTTGAAGAGCTCTCACTAATCTTCACTCACTCTC	1717
QY	3961	ATTATTAATGAAATCACTTGGGCTGTGGTCAACAGAAATGATTAATTTTAAATTCAGAA	4020
Db	1716	ATTATTAATGAAATCACTTGGGCTGTGGTCAACAGAAATGATTAATTTTAAATTCAGAA	1657
QY	4021	CCTTCATATTAAGTCATCTATATTTGCTAATATGACGAGAAAGCAACTCTTAACT	4080
Db	1656	CCTTCATATTAAGTCATCTATATTTGCTAATATGACGAGAAAGCAACTCTTAACT	1597
QY	4081	GCAATTAACAATCTAATTAATTAATAGTAAAGCAATCTTCCTTAAAGTTTAAATTCATTTG	4140
Db	1596	GCAATTAACAATCTAATTAATTAATAGTAAAGCAATCTTCCTTAAAGTTTAAATTCATTTG	1537
QY	4141	TGAGCAAGCTGTATTAATTTGGTGGGGCTCAAGCGCGGCTGTGTGAAATTCACAAAT	4200
Db	1536	TGAGCAAGCTGTATTAATTTGGTGGGGCTCAAGCGCGGCTGTGTGAAATTCACAAAT	1477
QY	4201	CAACAGATGTAGCGCTCTCGGCTTAAGTAAAGAAAGAAATGCAAGTTTAAATAGCT	4260
Db	1476	CAACAGATGTAGCGCTCTCGGCTTAAGTAAAGAAAGAAATGCAAGTTTAAATAGCT	1411

[illegible]

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Kelly, B., Kelly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J.,  
 Sanders, W., Saverly, G., Scherer, S., Scott, G., Shattman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajd, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K.,  
 Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, R., Warren, R., Mel, X., White, F.,  
 Williams, G., Wilson, R., Wlezyk, R., Wooden, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, S., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weise, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstock, G. and Gibbs, R.A.

Unpublished  
 2 (bases 1 to 160574)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (17-OCT-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 20, 2002 this sequence version replaced gi:24080629.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a BACfold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 Project information  
 Center project name: KBUG  
 Center clone name: CH230-436C20

----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 15133 bases at least Q40  
 Consensus quality: 15144 bases at least Q30  
 Consensus quality: 15429 bases at least Q20  
 Estimated insert size: 155447; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 Bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draai\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1 93077: contig of 93077 bp in length  
 93078 93177: gap of unknown length  
 93178 113865: contig of 20688 bp in length  
 113866 113965: gap of unknown length  
 113966 158783: contig of 44817 bp in length  
 158783 158882: gap of unknown length

FEATURES  
 source 158883 160574: contig of 1692 bp in length.

misc\_feature  
 1. 160574  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-436C20"  
 complement(5637..6655)  
 /note="clone boundary"

misc\_feature  
 end sequence:B203757"  
 complement(90372..91276)  
 /note="clone boundary"

misc\_feature  
 end sequence:B203758"  
 93178..96221  
 /note="wgs\_end\_extension"

## ORIGIN

Query Match 14.1%; Score 843.2; DB 2; Length 160574;  
 Best Local Similarity 70.3%; Pred. No. 3.9e-151;  
 Matches 1417; Conservative 0; Mismatches 488; Indels 111; Gaps 17;

QY	3981	GGCTGTGTGTCAGGAAATTGATTTTAAATTCAGAACTTCTATTAAGTCATCTA	4040
DB	65379	GGCCATGCTACAGGACATTAATTTCTCAATTTTAAATTTGCAATTCGATCGT	65320
QY	4041	TATTTGCTAATTCAGGAGGAAAGCAACTCTTTAATTCGCAATTAA-CAAATCTATA	4099
DB	65319	TATTTGCT-----CAAGAGAAAGAACTCAAACTCAATTCGCAATTCATATA	65265
QY	4100	TTAATTGCTAATTCAGGAAATTTCTTTAATTTTCAATTTTGTGAGCAAGCTTTGAT	4159
DB	65264	TTAATTGCTAATTCAGGAAATTTCTTTAATTTTCAATTTTGTGAGCAAGCTTTGAT	65205
QY	4160	TGGCTGGGCTCAGCGCCGCTTTTGTGAATTTTCAAGATTCAGAGATTCAGCGCT	4219
DB	65204	TGGCTGGGCTCAGCGCCGCTTTTGTGAATTTTCAAGATTCAGAGATTCAGCGCT	65145
QY	4220	CGGCTAATTCAGGAAAGGAAATTCGAAATTTTAAATTTGCTTCCATCTGGCT	4279
DB	65144	CAGCTGAGGAAAGGAAATTCGAAATTTTAAATTTGCTTCCATCTGGCT	65093
QY	4280	GAAAGCAATTCAGGAAATTTTAAATTCAGGAAATTCAGGAAATTCAGGAA	4339
DB	65092	GAAGCAATTCAGGAAATTTTAAATTCAGGAAATTCAGGAAATTCAGGAA	65034
QY	4340	ATGTCGAATTTCTCTGAAAGGCTTTAATTTGCTTCAACTTTGACATCTGATGTC	4395
DB	65033	ATGTCGAATTTCTCTGAAAGGCTTTAATTTGCTTCAACTTTGACATCTGATGTC	64983
QY	4400	ACCTATTACAGGTGTGTCTGTGACTAGGAGGAGTGAAGAGATTCGATCAATGT	4459
DB	64982	ACTATTGTCAGGTGTGTCTGTGACTAGGAGGAGTGAAGAGATTCGATCAATGT	64927
QY	4460	TAGTACCGCTTGAATTCAGGAGTGTGTTTTCCTGTTGAGGCTATTCGATCTG	4519
DB	64926	CAGGACATTTGAGGACAGG-----CTATTTTCCCTTGTGAGGCTATTCGATCTG	64874
QY	4520	ACCTTCTGAATTCATTTTCAATTTTCAATTCGAAATTCGAAATTCGAAATTCG	4579
DB	64873	ACCTTCTGAATTCATTTTCAATTTTCAATTCGAAATTCGAAATTCGAAATTCG	64814
QY	4590	CCATATTCAGAAAGAAATTAATTTTGTGTGACTTT-CCTGATTTTACACTGAT	4638
DB	64813	CTACATTCAGAAAGAAATTAATTTTGTGTGACTTT-CCTGATTTTACACTGAT	64754
QY	4639	TTGGAAATTCAGCAATTTTATGCTTCTTGAATAGTAGTCAAGTCAAGCAAC	4698
DB	64753	TTTCAATTCAGCAATTTTATGCTTCTTGAATAGTAGTCAAGTCAAGCAAC	64694



OY	4699	CAAAAACAGAAAACTGTAAAGACATTAAGATTAAGGTGAGGCCACCTGAGATTTAAA	4758
Db	64693	CAACTAAAGTAGAGGTAAAGATTAAGACACCTGATAGAGTTGAGGTCACTTAT	64633
OY	4759	TAAACTGAATAATTTTATTAACAGCAATTTGAAATAATTGTGCACTTCAGAAATTC	4818
Db	64693	TTTTCCATTGACGGGCTATCTGAATTAATCTCAATTAATTAATTCCAATTAATTAACAT	64574
OY	4819	TACATTAATTAATTAATTTCCAAATTTTAATCTTAAGAAATTAATATATATATGTA	4878
Db	64573	TAAAGAACATTAATTAAGAAACCAATTAACCTATTAATCTCTATTTTAATATCAATAAA	64514
OY	4879	GTACATGTCATCTGTCTTGAAGTGAATATTTAACTCAATAAGGTATTTTCTTTAT	4938
Db	64513	ATACATAATCTTACATCT---GTATGCAATAAAATCAACAAGGCTATTTTTTTAAT	64457
OY	4939	CCGGTAGACCAAGCCTCTAAGGGAGATGAAGGAATACTCTTCTCT-----	4988
Db	64456	TGAGAGAAATAATTTCTAATTAATAGCAATCTCTCTCTCTCTCTCTCTCTCTCTCT	64397
OY	4989	-----TACGTAGAGAGAGTGAATTTAC	5015
Db	64396	CT	64337
OY	5016	TTAATATATTAACAGAAATTTCCCTGCTCTTTGTCTATTTGATTTGTGACCAACAAGCG	5075
Db	64336	TTAAATATATTAACAGAAATTTCCCTGCTCTTTGTCTATTTGATTTGTGACCAACAAGCG	64277
OY	5076	GTGTGCTGAAGGGAATCTGAAGGCCGGAGAGGCAAGCAATATCAATGAATAACAAAC	5135
Db	64276	GTGTGCTGAAGGAGGCTGAAGGCCGGAGGTGTGGGAG-----GAAAGAAAGGGG	64226
OY	5136	AAACAAAAATTCCTTAACAGCTCTCAACAAATTTTGAAGCCAGAAATATGACAGAA	5195
Db	64225	AAAAAATCTCTCTAAGGCTCTCTGCAACAGCTTACCAACAGCACTGACAGAG	64166
OY	5196	TCTCTAATCAAAACCAAGTATCCAGATTAACAAGAGTGTATGTAGCTGACAGGGTGA	5255
Db	64165	CTCTCAATCAATCAAGATTAAGATTAACAAGAGTGTATGTAGCTGACAGGGGGA	64106
OY	5256	CACTCATCACTCAAGTTCAATTCAAAAGTCCAGGCTGCTGAATAATTAATCTGATGCA	5315
Db	64105	CATTATTAATCACT-----GACAAAGCTCGCTCCCGATTAATCACTGATGCA	64053
OY	5316	TTCTATCCAGATTCATCAAGACAGAGATCAAGAGTTCAAGATGCTCACTCAAA	5375
Db	64052	TTCTATCCAGATTCATCAAGACAGAGATCAAGAGTTCAAGATGCTCACTCAAA	63998
OY	5376	TTGCAACAACAAGTGTGGGCTAATATAGCTGAAGCACTGAAAGCTGAGAGAGCGGA	5435
Db	63997	----TGCCAAAGTGTGGCACTATATGTCAGAGGCTTAATCTGTGGCAACAGGA	63943
OY	5436	AGAACACAGTAAAGAGAGATGCCAGCTGTAAAGATCAAGTGTTAAGAGTTTATGTC	5495
Db	63942	AGAACACAGTAAAGAGAGATGCCAGCTGTAAAGATCAAGTGTTAAGAGTTTATGTC	63883
OY	5496	AATGATGATCTCAATTTGGCTAAATAATCAAGAAACGCTCGGCTCTTTGCAAAATATG	5555
Db	63882	ACTTCAAGACGCTATTTGGCTGAAGAGAAATGCTCGGCTCTTGAAGAAATCTGTGA	63823
OY	5556	AAGGAGAGAGGCTCAAACTTCTAATGTCTGATTAAGATTAAGCTTATAGCT	5615
Db	63822	AAGGAGAGAGGCTCAAACTTCTAATGTCTGATTAAGATTAAGCTTATAGCT	63763
OY	5616	CCCGGCTTATCTATATATTAACAAGATTTGTGTATATTTTATTAATATGTTCTCGG	5675
Db	63762	CCCGGCTTATCTATATATTAACAAGATTTGTGTATATTTTATTAATATGTTCTCCCA	63703
OY	5676	TTCTGTATATCAAGACAGTTGAAGAAATTTGAAATCTCAAGGTGTGCACGC	5735
Db	63702	TCGTGTATATCAAGAGGTGTGAAGAAATTTTGAATCTCAAGGTGTGCACGC	63643

QY 5733 TACAGTATCGCCCTCGACCCCTTAATCATCAGGGGTGGTTGAATGTGACATPACTCTCTC 5795

Db 63642 TACAGTATCTGCCCTCGACCCCTTAATCATCAGGGGTGGTTGAATGAGCTACGCTCTGTC 635833

QY 5796 AAAACTTAATGGAAGTCCCTGTGTGATTAATGGAATGTGACCTGTGTGACAAAGACATTT 5855

Db 63582 AAAACTGAAGCTGGTGTCC - TGTGTAATCTGAGGTGTCTCTGCTTTACTA -- AAGAGC 635282

QY 5856 CCTCCTTGTATTAGTCAGACAGAGTGATTAATTTTGGACTTTCTGTGACTTAAGAGTGC 5915

Db 63525 CCTCCTTGTATTAGTGGCCATTAAGAACTACTTGTGACTTTGTGTGACTTAAGAGTGC 63466

QY 5916 TGTGACATATTTCTGAATGTCTTTTTGTGTGAT 5951

Db 63465 TATTGACATCTTTTCTAAGTAACTTAATACTTAAT 63430

RESULT	8
AC113784/C	
LOCUS	284005 bp DNA linear HTG 19-NOV-2002
DEFINITION	Rattus norvegicus clone CH230-96F2, WORKING DRAFT SEQUENCE, 3
ACCESION	AC113784
VERSION	AC113784.5
KEYWORDS	HTG; HTGS_PASB1; HTGS_DRAFT; HTGS_FULLTOF.
SOURCE	Rattus norvegicus
ORGANISM	Rattus norvegicus (Norway rat)

## REFERENCE AUTHORS

1 (bases 1 to 284005)

REFERENCE  
AUTHORS

Murphy, D. Marie, Melker, M. Lee, Abbanzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alb Brooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaratatke, D., Barber, M., Barnstead, M., Barnhead, P., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Centar, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dedertich, D., Delgado, O., Denison, S., Deramo, C., Ding, F., Dinh, H., Dlyak, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durkin, K., Duval, B., Eaves, K., Egan, A., Encotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabris, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georgette, S., Geer, K., Gail, R., Grady, M., Guerra, M., Guevara, W., Guarratone, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Hulys, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacobs, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kwie, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, Y., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensbuewa, L., Louiseged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratna, M., Mamoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Manganey, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minje, E., Montemeyor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundesa, M., Murphy, M., Naitir, L., Nankariva, C., Nel, D., Newton, N., Nguyen, N., Norris, S., Nuckelamch, O., Okonum, G., Oltarnpasagon, A., Pal, S., Parks, K., Paternack, S., Pali, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Polndexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, J., Scherer, S., Scott, G., Sheltman, S., Shen, H., Shetty, J., Shivatspeyn, A., Sisson, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sovelle, R., Soza, J., Steinkle, M., Strong, R., Sutton, A., Syatek, A., Tabot, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K.,



TITLE	Direct Submission
JOURNAL REFERENCE	Unpublished 2 (bases 1 to 284005)
AUTHORS	Morley,K.C.
TITLE	Submitted (05-MAR-2002) Human Genome Sequencing Center, Department of Molecular Biology and Genetics, Baylor College of Medicine, One
JOURNAL	Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Mang,O., Mang,S., Warren,J., Warren,R., Wei,X., White,P., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,T., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,Z., Zhou,S., Zhao,S., Dunn,D., Von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

3 (bases 1 to 284005)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular Biology and Genetics, Baylor College of Medicine, One

**COMMENT**

On Nov 19, 2002 this sequence version replaced g1:23195097. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.com/bmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a contig-scaffold). Within each contig-scaffold, individual sequence contigs are ordered and oriented and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu>  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: GSCN  
Center clone name: CH230-9882

```
Assembly program: Phrap, version 0.990329
Consensus quality: 245691 bases at least Q40
Consensus quality: 252542 bases at least Q30
Consensus quality: 251124 bases at least Q20
Estimated insert size: 252183; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
```

- NOTE: Estimated insert size may differ from sequence length. (see [http://www.birc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.birc.bcm.tmc.edu/docs/Genbank_draft_data.html))
- NOTE: This sequence may represent more than one clone.
- NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
- This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

## FEATURES

```
source 1..384005
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-98P2"
misc_feature 1..1035
```

```

/note="wgs_end_extension
clone_end:T7"
misc_feature
7353..8173
/note="clone_boundary
clone_end:T7
site:
end_sequence:RH306729"
9665..11350
/note="wgs_contig"
14183..15278
/note="wgs_contig"
15329..16533
/note="wgs_contig"
complement(261454..264092)
/note="clone_boundary
clone_end:Sp6
site:
end_sequence:RH306731"
274560..275276
/note="wgs_end_extension
clone_end:Sp6"

```

Query Match	14.1%	Score 843.2	DB 2	Length 284005
Best Local Similarity	70.3%	Pred. No. 3,7e+151		
Matches 1417	Conservative	0	Mismatches 488	Indels 111
			Gaps	17
Qy	3981	GGCTGTGTCACAGGAAATTGATTTATTTTAAATTCGAAACCTTCATTTAGTCATATCA		40404
Db	72189	GGCCATATGCTAACAGACATTTGTAATTCGAAATTAAGAACTTGCAATCGGCTATCTG		72131
Qy	4041	TATTTGCTATATGACAGGAGAGAAAGCCAACTCTTAACTGCAATTA-CAATCTATTA		40995
Db	72129	TATTTGCT-----CAGAGAAAGAGTCAAACTTCATTTCAATGACAAATCATATA		72077
Qy	4100	TTAATAGTGAAGCAATCTCCCTTAAGTTTACATTTGTGAGCAAGCTGTGAT		41575
Db	72074	TTAATTTGTCATATGATCTTTTGCTTCAAGCTTCACATTTTATGAAAAGCTCTGAT		72070
Qy	4160	TGCTTGCGGCTCAAGCCCGGCTGTGTGAAATTCACAAATTCACAAATTAAGCGCT		42151
Db	72014	TGCTTGCGGCTCAAGCCCGGCTGTGTGAAATTCACAAATTCACAAATTAAGCGCT		71995
Qy	4220	CGGCTTAAGTAAAGGAAGAAATGTCAAGTTTAAATAGCTTCTCCCTTCATCTGCT		42757
Db	71954	CAGTTCGAGGAAGGAAGAAATGTCAAGTTTAAATAGCTTCTC-----CGGCT		71900
Qy	4280	GAAGCAACAAATTAATTTTATGAAACATTTGATGATTTACTTAACAGGAA		43355
Db	71902	AGAAACAACAAATTAACATCTTTAG-AAACATTTGAAACGAATTTCTTAACAGAA		71865
Qy	4340	ATGTCAAAATTTCTTGAAAGGCTTAAATGTCTCAACTTGCATCTACATGTC		43995
Db	71843	ATGTCAAAATTTCT-----CTCAACTGTGCTCAACACTTGAATTTTCTGATGTC		71777
Qy	4400	ACCTATTTACAGGTGTGTCTCTGTGACTGAAGGAGGAGTGAATCTACAGAT		44555
Db	71792	ACATGTTTCAAGGTGT-----CTATGACTACGGCAGTGAATGAG-AGATATGTACTTGAATGT		71777
Qy	4460	TATGACCCGTTATATCAAGAGTGTTTTTTTTCCCGCTGTGAGTCTATCTAAGT		45155
Db	71736	CAGGACACTTGGCGACACG-----CTCATTTCCCTCTTGTGACTCATTTCTAATTT		71665
Qy	4520	AGCTTCGATCATTTTGTTCATTTTCGAAATTCGACAAACAGGATAGTTTACAG		45775
Db	71683	AGCTTCGATCATTTTGTTCATTTTCGAAATTCGACAAACAGGATAGTTTACAG		71665
Qy	4580	CCATATTCAGAAAGAAATTAATTTTGTGTGACCTT-CTCATTTACATGAT		46355
Db	71623	CTCATTTCAAAAGAAACAAATTAATTCATGACGCTTGTGCTCATTAATTAATTAAT		71555
Qy	4639	TTGGAAATATATGAAACAAATTTTATGCTTCTTTCGAAATGTGCAAGTCAAGCAAAAC		46955

[illegible][illegible]

RESULT 9	AC126304/c	LOCUS	AC126304.6	DEFINITION	Rattus norvegicus clone CH230-244017, *** SEQUENCING IN PROGRESS	20313 bp	DNA	linear	HTG 20-NOV-2002
ACCESSION	AC126304	VERSION	AC126304.6	KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.	GI:25136584			
SOURCE	Rattus norvegicus	ORGANISM	Rattus norvegicus		(Norway rat)				

## REFERENCE AUTHORS

1 (pages 1 to 204313)

Munzy, D., Merte, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alb Brooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Banatvala, K., Barber, M., Barnstead, M., Benahmed, F., Bielawski, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burck, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cessari, H., Chen, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Andrade, C., Dedrich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Dlyva, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Secotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganter, R., Garcia, A., Garner, T., Garza, H., Gebregiorgis, E., Geer, K., Giller, R., Grady, M., Guerra, I., Guevara, W., Gunaratne, P., Halaland, M., Hamill, C., Hamilton, C., Hamilton, K., Harvey, I., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huily, S., Hume, J., Idelblut, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Joivet, A., Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, J., Kovcs, C., Krefft, C. L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J., Lornshwara, L., Louisedge, H., Lozdo, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Manjuma, A., Manum, B., Mapus, P., Martin, K., Martin, R., Martinez, E., Mashinay, S., McLeod, M. P., McKell, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwachilemeh, O., Okunolu, G., Olarunpusgoun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polndexter, A., Popovic, D., Plums, E., Pu, L., L., Puazo, M., Quiror, J., Rachlin, E., Reeves, K., Reider, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rivers, C., Rodkey, T., Rojts, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savory, G., Scherer, S., Scott, G., Schatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smalls, D., Sneed, J., Sdergren, E., Song, X., -Z., Sorrelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, J., Tabot, P., Taylor, C.,

BEST AVAILABLE COPY

Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wlezyk, R., Wood, H., Morley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

**TITLE**  
JOURNAL  
Unpublished

**REFERENCE**  
2 (bases 1 to 204313)

**AUTHORS**  
Morley, K.C.

**TITLE**  
Submitted (05-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

**JOURNAL**  
3 (bases 1 to 204313)

**REFERENCE**  
Rat Genome Sequencing Consortium.

**AUTHORS**  
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

**COMMENT**  
On Nov 20, 2002 this sequence version replaced gi:2397953. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

**Genome Center**  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

**Project Information**  
Center project name: GZGP  
Center clone name: CH230-244D17

**Summary Statistics**  
Assembly program: Phrap; version 0.990329  
Consensus quality: 190230 bases at least Q40  
Consensus quality: 19082 bases at least Q30  
Consensus quality: 194928 bases at least Q20  
Estimated insert size: 197682; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

**NOTE:** Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
**NOTE:** This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

**FEATURES**  
**Source**

1	71914	contig of 71914 bp in length
71915	72014	gap of unknown length
72015	203206	contig of 131192 bp in length
203207	203306	gap of unknown length
203307	204313	contig of 1007 bp in length.

**Location/Qualifiers**

1..204313	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10116"	
/clone="CH230-244D17"	
complement(4101..4830)	

**misc\_feature**

**misc\_feature**  
/note="clone boundary  
clone\_end:77  
site:"  
end sequence:B2158614"  
complement(69055..63921)  
/note="clone boundary  
clone\_end:5p6  
site:"  
end sequence:B2158615"  
72015..73525  
/note="wgs\_end\_extension  
clone\_end:5p6"  
19105F..192396  
/note="wgs\_end\_extension  
clone\_end:5p6"

**misc\_feature**  
/note="wgs\_end\_extension  
clone\_end:5p6"

**ORIGIN**

Query Match 14.0%; Score 835.8; DB 2; Length 204313;  
Best Local Similarity 70.0%; Pred. No. 9.9e-150;  
Matches 1412; Conservative 0; Mismatches 493; Indels 111; Gaps 17;

QY	3981	GGCTGGTCAGAGAAATGATATTTTATTTTGAAGACCTCTATTAAGTCATCA	4040
DB	63951	GGCTATGCTAGAGACATTAATTTCTCAATTTAGAACTTGCATCGGATCTG	63892
QY	4041	TATTTGCTAATAGCAGGAGAAAGCAACTTTTAATGCAATTA-CAATCTATA	4099
DB	63891	TATTTGCT-----CAGAGAGAAAGATCAACCTTCAATTCATAGAGAAATCTATA	63837
QY	4100	TTAATAGTTAGCAATCTTCCCTTATAGTTTATTCATTTTGGAGCAAGCTTGAT	4159
DB	63836	TTAATAGTTAGCAATCTTCTGCTTCAAGCTTCAATTTTGGAGCAAGCTTGAT	63777
QY	4160	TGGCTGGGCTCAGCGCGGCTTTCGAAATTCAGAAATTCAGAGTTTACCGCT	4219
DB	63776	TGGCTGGGCTCAGCGCGGCTTTCGAAATTCAGAAATTCAGAGTTTACCGCT	63717
QY	4220	CGGCTAAGTAAAGAGAAATTCAGATTTAATAGCTTCCCTTCATCTGCT	4279
DB	63716	CAGCTAGAGAAAGAGAAATTCAGATTTAATAGCTTCCCTTCATCTGCT	63665
QY	4280	GAAACCAAAATTAATTTTATTAAGAAACAAATTTAGTTAGTTACTTACAGGAA	4339
DB	63664	GAACCAAAATTAATTTTATTAAGAAACAAATTTAGTTAGTTACTTACAGGAA	63606
QY	4340	ATGTCAAATTTCTCTAAGAGGCTTATGATGTCACAACTTTCATCTAGATTC	4399
DB	63605	ATGTCAAATTTCTCTAAGAGGCTTATGATGTCACAACTTTCATCTAGATTC	63555
QY	4400	ACCTATTACAGGTGATCTGTAAGAGGAGTAAAGGAAATGTAATCAATCAAT	4459
DB	63554	ACTATTTCAGAGGTGATCTGTAAGAGGAGTAAAGGAAATGTAATCAATCAAT	63499
QY	4460	TAGTACCGTTAATACAGAGATGTTTTTCCCGCTGAGATCTATCTAATCAG	4519
DB	63498	CAGGACATTTGGCAGAGG-----CTCATTTCCCTTGGACTCATCTAATTT	63446
QY	4520	AGCTTGAATCATTTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCA	4579
DB	63445	AGCTTGAATCATTTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCA	63386
QY	4580	CCATATTCAAGAAAGAAATTAATTTTGTGTGAGCTTT-CTGATTTACACTGAT	4638
DB	63385	CTACATTTCAAGAAAGAAATTAATTTTGTGTGAGCTTTCTGCTGATTTATTAAT	63326
QY	4639	TTGGGAATATTAACAATTTTATGCTTTCTTCAAGATGAGCTCAATCAAGCAAA	4698
DB	63325	TTTCTACTATTAACAATTTTATGCTTTCTTCAAGATGAGCTCAATCAAGCAAA	63266
QY	4699	CAAAACAGCAAAATCTGTAAGACATTAAGATGAGAGCCGACTGAGATTAATA	4758
DB	63265	CAACTAAAGTGAAGATTAAGATTAAGACAGACTGATGATGATGATGATCACTTAT	63206



TITLE  
 JOURNAL  
 COMMENT  
 Submitted (04-OCT-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Oct 4, 2003 this sequence version replaced g1:34482079.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L23503  
 Center clone name: 186\_N\_8

FEATURES  
 source  
 1. 214853  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="16"  
 /map="16"  
 /clone="RP23-186N8"  
 /clone\_lib="RPC1-23 Female Mouse BAC"  
 1. 51  
 /note="wgs and extension  
 clone end:5p6"  
 52. 57  
 /note="clone boundary  
 clone end:5p6  
 site:ECORI"  
 1042. 1080  
 /rpt\_family="(CATATA)n"  
 1179. 1367  
 /rpt\_family="B3"  
 2436. 2578  
 /rpt\_family="LM4C"  
 2651. 2737  
 /rpt\_family="LM4C"  
 complement(2850..3351)  
 /rpt\_family="RMKR19B"  
 repeat\_region  
 complement(3350..3516)  
 /rpt\_family="B3"  
 repeat\_region  
 3527. 3891  
 /rpt\_family="LM4C"  
 4170. 4563  
 /rpt\_family="LM4C"  
 4873. 4952  
 /rpt\_family="B1\_MM"  
 4968. 5018  
 /rpt\_family="(TGAA)n"  
 5056. 5415  
 /rpt\_family="MTD"  
 5554. 5810  
 /rpt\_family="LM4C"  
 6353. 6500  
 complement(6731..7895)  
 /rpt\_family="Lx8"  
 7896. 7919  
 /rpt\_family="(TTTTC)n"  
 complement(7920..8304)  
 /rpt\_family="Lx7"  
 repeat\_region  
 complement(8305..9900)  
 /rpt\_family="Lx4"  
 repeat\_region  
 complement(9899..10500)

/rpt\_family="Lx4"  
 repeat\_region  
 10516. 10639  
 /rpt\_family="L1"  
 repeat\_region  
 10634. 11419  
 /rpt\_family="L1"  
 11420. 11615  
 /rpt\_family="B2\_Mm2"  
 11616. 13898  
 /rpt\_family="L1"  
 complement(13899..14142)  
 /rpt\_family="LTRIS\_MM"  
 14447. 17150  
 /rpt\_family="Lx8"  
 17251. 17303  
 /rpt\_family="(TGC)n"  
 17523. 17580  
 /rpt\_family="AT\_rich"  
 17590. 17667  
 /rpt\_family="(GAAA)n"  
 17669. 18434  
 /rpt\_family="Lx8"  
 18558. 18593  
 /rpt\_family="AT\_rich"  
 18855. 19268  
 /rpt\_family="BGLIT"  
 21016. 21156  
 /rpt\_family="B1\_MM"  
 21157. 21199  
 /rpt\_family="(CAA)n"  
 complement(21371..21767)  
 /rpt\_family="LIME"  
 complement(21768..21878)  
 /rpt\_family="L1"  
 21869. 22524  
 /rpt\_family="L1\_MM"  
 22533. 22555  
 /rpt\_family="(A)n"  
 complement(22556..23200)  
 /rpt\_family="L1"  
 complement(25514..25708)  
 /rpt\_family="B3"  
 26244. 26293  
 /rpt\_family="MT2B"  
 26294. 26369  
 /rpt\_family="B1\_MM"  
 26395. 26369  
 /rpt\_family="ID\_B1"  
 26370. 26532  
 /rpt\_family="(GAAA)n"  
 26533. 26553  
 /rpt\_family="ID\_B1"  
 26674. 26921  
 /note="single clone coverage"  
 26678. 26712  
 /rpt\_family="(TTTG)n"  
 complement(26715..26836)  
 /rpt\_family="B1\_MM"  
 26898. 26945  
 /rpt\_family="(TC)n"  
 26945. 26989  
 /rpt\_family="(TG)n"  
 complement(26992..27264)  
 /rpt\_family="B4A"  
 30432. 30458  
 /rpt\_family="(T)n"  
 30639. 30668  
 /rpt\_family="(CCCCCT)n"  
 complement(31137..31532)

Query Match 14.64; Score 835.4; DB 10; Length 214853;  
 Best Local Similarity 70.24; Pred. No. 1.2e-149;  
 Matches 1439; Conservative 0; Mismatches 471; Indels 141; Gaps 18;

QY 3981 GCCTGCTGACAGAAATGATTTATTTATTCAGAACTCTTAATAGCATCTA 4040  
 DB 51809 GTCCATGCGCATAGCAAAATGATTTATTCATTTAAGAACTTAATATGAGGATCTG 51868  
 QY 4041 TATTTGCTAATAGCAAGAAAGAAAGCAATCTTTAACTGCATTAACAATCTAAT 4100  
 DB 51869 TATTTGCTCAAA--GGAAGAAAGCAAAACG--TTGATTCATTAAGCAATCTAAT 51923  
 QY 4101 TAAATAGTAAGCAATCTTCCCTTTAGTT--TACATTTTGTGAGCAAGCTGTGAT 4158  
 DB 51924 TAAATAGTAAGCAATCTTGTCTCAAGTTAACATTTTTTTTGAAGCAAGCTGTGAT 51983  
 QY 4159 TTGGTGGGCTGAGCGCGCGCTGTTGAAATTCACAAATTCAGATGAGCGCTG 4218  
 DB 51984 TTGGTGGGCGCAAGCGCGCAATGAGATGAGATTCACAAATTCAGATGAGCGCG 52043  
 QY 4219 TCGGCTAAGTAAGAAAGAAAGTAATGCAATTAATAGCTTCCCTCCATCTGAGC 4278  
 DB 52044 TCAATCTGAGTTAAGAAAGAAAGTAATGCAATTAATAGCTTCC--CTGCGC 52095  
 QY 4279 TGAAGCAACAATTAATATTTTATGAAACATTTTGAATTAATTAAGTCTTAAGGGA 4338  
 DB 52096 TGAAGCAACAATTAATATTTTATG--AAGCAATTTTGAAGCAATTTCTTAAGGGA 52154  
 QY 4339 AATGTCAAATTTCTGTAAGAGGCTTTAATTTGCTCAAACTTTGACATCTAAGAT 4398  
 DB 52155 AATGTCAGATTTCTC-----TTGAGCTGCTCAAACTTTGACATTTCTAATGT 52205  
 QY 4399 CACCTATTTAGAGGTGCTGCTGTAAGGAGGAGGAGGAGATGCACTCAACATG 4458  
 DB 52206 CACTAGTTTGCAGGTG--CTATGCTGTGAGCAAGAAAG--ACATATG--ACGTAACTG 52261  
 QY 4459 TTAGTACCGCTAGATCAAGAGGTGTTTTTTTTTCCCGTGTGAGTATCTAATCT 4518  
 DB 52262 TCAGTACACTGTGGCAACAG-----CTTACTGCTGCTGCTGAGCTCATTTCTAAT 52314  
 QY 4519 GAGCTTGCATCTAATTTGATTCATTTCCAAATTCACAAACCAAGATTAAGTACG 4578  
 DB 52315 TAGCTTGCATCTAATTTGATTCATTTCCAAATTCACAAACCAAGATTAAGTACG 52374  
 QY 4579 CCCATTTCCAGAAAGAAATTAATTTTGTGTAGATCTTCTGATATTAACATGAT 4638  
 DB 52375 CCTACATTCAGAAAGAAACAAATTCATTTCACTGAGACCTTCTGATATTAAT 52434  
 QY 4639 TTGGGATATATGAACAATTTATGTTTCTTGGAGTGAAGTCAAGCAAGCAAAAC 4698  
 DB 52435 TCTTTACATATTAAGAT-----GTCAATTTTATGTAAGTGAAGTGAAGCAAGCT 52488  
 QY 4699 CAAAAACGCAAAACCTTAAGCATTAAGAAATGAAGGAGCGAGCTGAAGATTA 4758  
 DB 52489 AAAAAAGTTAAA--GTGAGGCAATTAAGATGAAGCAAGATGAAGATTTGAAAT 52545  
 QY 4759 TAAACTAGAAATTTTATTAAGCAAGCAATTCAAATTAATTTGCACTTCAGATATTC 4818  
 DB 52546 AACTTATTTTCCCATTCAGAGCGCATTAATAATTTCTAATAAATATGAAAACA 52605  
 QY 4819 TACAATAATATATATTTCCAAATTTTAATTTTAAAGAAATTAATTAATTAATGTA 4878  
 DB 52606 AATAAACATTAAGAAACATTTGAAAGAAATTAATTAATTTTATTAATTAATCAATAAA 52665  
 QY 4879 GTACATGTGCAATGTTTGAAGTGAAGATTTTACTCAATTAAGGATTAATTTCTTTAT 4938  
 DB 52666 ATCCATTAATTAATATG--TATG--TAAATAAATAATTAATTAATTAATTTCTTTAT 52724  
 QY 4939 CCGGTCAAGCAAGCTTCTTAAGGAGATGTAAGGAGATATCTCTTCTCTC----- 4988  
 DB 52725 TGGGACAGAAACATTTCTAATAATTAATTAATTAATTAATTTCTCTCTCTCTCT 52784  
 QY 4989 ----- 4988  
 DB 52785 CT 52844  
 QY 4989 -TAGCTGAGAGAGAGTGAATTTCAATTAATTAATTAATTAATTAATTAATTTCTCTCTCTG 5047

DB 52845 CTGTTGAAATGAAGATGATGTTCTAGTTAATTAATTAATTAATTAATTTCTCTCTTTG 52904  
 QY 5048 CTATTTAGATTTGACACACACAGGCGGTGGCTGAAGGAACTGAAGGCGGAG 5107  
 DB 52905 CTATTTAGATTTGACACACACAGGCGGTGGCTGAAGGAACTGAAGGCGGAG 52964  
 QY 5108 GGAAGCAATATGATGAATAAACAACAACAACAATCTTCCCTTGAAGCTCTTCAAAAC 5167  
 DB 52965 GTGGG-----CAGGAGAAAGCGAAGAAAAAAGTCTCTAAGCTCTTCCAAAC 53019  
 QY 5168 ATTTTACCCCAAGAAATATGACACAGAAATCTCAATTAACCAAGCTATCCAGATCAAG 5227  
 DB 53020 ATTTGACACAGCAAGCAATGACAGAGCTTCAATTAAGTCAAGTAATTAATTAATTAAG 53079  
 QY 5228 AATGTTATGATGACAGAGGAGTGAACATCACTACGCTGAGTTCAAAAGTCC 5287  
 DB 53080 AATGTTATGATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 53132  
 QY 5288 AGCTGCTGAATTAATTAATCTGATGCAATTCAGCAAGCTCAATGACAGAGATCA 5347  
 DB 53133 TGGCTTCTGATTAATTAATCTGATGCAATTCAGCAAGCTCAATGACAGAGATCA 53192  
 QY 5348 GAATTCAGAGATGCTCTCACTCCAAATTCGCAACAAAGTGTGCTACTAATCTCA 5407  
 DB 53193 GAATTCAGAGATGCTCTCACTCCAAATTCGCAACAAAGTGTGCTACTAATCTCA 53243  
 QY 5408 AGCACTGAGAGCGCTGAG 5467  
 DB 53244 AGCGCTCAAACTGTGAGAG--GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 53302  
 QY 5468 ACAATGAGCTTATTAATTAATTTTATGATTAATTAATTAATTAATTAATTAATTAATTA 5527  
 DB 53303 ACAATGAGCTTATTAATTAATTTTATGATTAATTAATTAATTAATTAATTAATTAATTA 53362  
 QY 5528 AGCTGCGCTCTTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5587  
 DB 53363 AGCGCGCGCTCTTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 53422  
 QY 5588 TACATTTGACCTATGCTTTTATGCTCCGCTTTTATCTAATTAATTAATTAATTAATTA 5647  
 DB 53423 TGGCATTTGACCTATGCTTTTATGCTCCGCTTTTATCTAATTAATTAATTAATTAATTA 53482  
 QY 5648 TGTGTAATTTTATTAATTAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5707  
 DB 53483 TGTGTAATTTTATTAATTAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 53542  
 QY 5708 AATTTGAACTTACCGGTGACCACTTACAGTACTGCGCTGACCTTACATCAAGCG 5767  
 DB 53543 AATTTGAACTTACCGGTGACCACTTACAGTACTGCGCTGACCTTACATCAAGCG 53602  
 QY 5768 TCAATTTGAATGACATTAATTTCTGCAAACTTAATTAAGAGGCTGCTGATTAATGA 5827  
 DB 53603 TCAATTTGAATGACATTAATTTCTGCAAACTTAATTAAGAGGCTGCTGATTAATGA 53662  
 QY 5828 ATGCTGACCTGCTGACAAAGCAATTTCTCTTTTATGACGACAGATATTAAT 5887  
 DB 53663 GTGTTTCTGCTGATG--TGAAGAGCGCTCTGTTTATGAGGAGCAATTAATTAATTAAT 53720  
 QY 5888 TTGACCTTCTGCTGACCTTAATGAGTCTGCTGACATATTTTCAATGCTTTTGGT 5947  
 DB 53721 TTGACCTTCTGACCTTAATGAGTCTGCTGACATATTTTCAATGCTTTTACT 53783  
 QY 5948 TCAATTTGA 5958  
 DB 53781 TTAATTTTGA 53791

RESULT 11  
 HSP63G01  
 LOCUS HSP63G01 825 bp DNA linear P01 04-JAN-2001  
 DEFINITION Homo sapiens P63 protein (P63) gene, exon 1.  
 ACCESSION AF124528















## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McChesron, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

## SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Prengen, E., Tatem, M., Catanesi, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.choir.org>

VECTOR: pBAC3.6

## NEIGHBORING SEQUENCE INFORMATION:

This sequence is not the entire insert of the clone. This clone is overlapped by AC11754 and AC09656.

There is a simple sequence repeat from base 4648 to 46739. This region does not meet required finishing standards.

## FEATURES

## SOURCE

```
1..159249 Homo sapiens
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="2"
  /map="2"
  /clone="RP11-733G6"
  /clone_lib="RPCI-11"
  102..178
    /rpt_family="LI"
  178..571
    /rpt_family="LI"
  571..626
    /rpt_family="MALR"
  626..942
    /rpt_family="LI"
  942..1024
    /rpt_family="LI"
  1024..1916
    /rpt_family="LI"
  1916..2217
    /rpt_family="LI"
  2217..2499
    /rpt_family="ALU"
  2499..3481
    /rpt_family="LI"
  3481..3782
    /rpt_family="LI"
  3782..3929
    /rpt_family="ALU"
  3929..4251
    /rpt_family="LI"
  4251..4764
    /rpt_family="ALU"
  4764..4986
    /rpt_family="LI"
  4986..5023
    /rpt_family="LI"
  5023..5176
    /rpt_family="(TA)n"
  5176..5236
    /rpt_family="LI"
  5236..5404
    /rpt_family="LI"
  5404..5520
    /rpt_family="LI"
  5520..5705
    /rpt_family="LI"
  5705..5777
    /rpt_family="LI"
  5777..5782
    /rpt_family="AT-rich"
  5782..6450
    /rpt_family="LI"
  6450..6474
    /rpt_family="LI"
  6474..6998
    /rpt_family="(TG)n"
  6998..7244
    /rpt_family="(TG)n"
```

```
repeat_region /rpt_family="MIR"
7948..8299 /rpt_family="MER2_type"
8299..8688 /rpt_family="MER2_type"
8688..9128 /rpt_family="MER2_type"
9128..9299 /rpt_family="MER2_type"
9299..9510 /rpt_family="MER2_type"
9510..9711 /rpt_family="ERVL"
9711..9717 /rpt_family="ERVL"
9717..10011 /rpt_family="ERVL"
10011..10275 /rpt_family="ALU"
10275..10313 /rpt_family="AT-rich"
10313..10375 /rpt_family="AT-rich"
10375..10528 /rpt_family="(TA)n"
10528..10529 /rpt_family="(TA)n"
10529..10659 /rpt_family="(TATAA)n"
10659..11412 /rpt_family="ERVL"
11412..11440 /rpt_family="ERVL"
11440..11544 /rpt_family="ERVL"
11544..11569 /rpt_family="ERVL"
11569..11927 /rpt_family="MALR"
11927..12495 /rpt_family="MALR"
12495..12856 /rpt_family="ERVL"
12856..13041 /rpt_family="ERVL"
13041..13197 /rpt_family="ERVL"
13197..14809 /rpt_family="LI"
14809..14829 /rpt_family="LI"
14829..14876 /rpt_family="LI"
14876..14898 /rpt_family="MIR"
14898..15187 /rpt_family="MIR"
15187..15243 /rpt_family="MER1_type"
15243..15523 /rpt_family="MER1_type"
15523..15728 /rpt_family="ALU"
15728..15999 /rpt_family="ALU"
15999..16186 /rpt_family="ALU"
16186..16418 /rpt_family="MALR"
16418..16727 /rpt_family="MALR"
16727..17022 /rpt_family="ALU"
17022..17166 /rpt_family="ALU"
17166..17468 /rpt_family="ALU"
17468..17670 /rpt_family="ALU"
17670..17868 /rpt_family="ALU"
17868..18330 /rpt_family="MIR"
18330..18674 /rpt_family="MIR"
18674..19048 /rpt_family="MIR"
19048..19069 /rpt_family="MIR"
19069..20886 /rpt_family="AT-rich"
20886..21207 /rpt_family="AT-rich"
21207..21379 /rpt_family="CR1"
21379..22209 /rpt_family="CR1"
22209..22342 /rpt_family="LI"
22342..22558 /rpt_family="LI"
22558..22572 /rpt_family="MIR"
22572..22872 /rpt_family="MIR"
22872..23219 /rpt_family="CR1"
23219..23264 /rpt_family="CR1"
```

Query Match 3.6%; Score 212.4; DB 9; Length 159249;

Best Local Similarity 63.3%; Pred. No. 1.2e-30;

Matches 413; Conservative 0; Mismatches 221; Indels 18; Gaps 5;

```
QY 3233 TGCCTGCTCATGCTGCTATATCCCAACACTTGGCAAGCCGCTCGGAGAGATAGCTTACG 3352
DB 85766 TGCCTGCTCATGCTGCTATATCCCAACACTTGGCAAGCCGCTCGGAGAGATAGCTTACG 85707
QY 3353 TCACGAGTTGAGACCACTGAGGCAACACAGCAAGACCCATATCTAATAAACAATA 3412
DB 85706 CTCGAGCATTTGAGACCACTGAGGCAACACAGCAAGACCCATATCTAATAAACAATA 85647
QY 3413 CAAACAAATATACCTGGGTATGCTGCTACCTGATGCAAGGTACACAGAAAGCT 3472
```

Db	85646	-----AAATTAAGCTGGGTGTGTGTGTACATGCTGTAGTCCAGGTACTCAGAGGCT	85593
Qy	3473	GAGGGAAAGATTAACCTTAGAGCCCAAGAGGTTGAGGCGGAGTATCATGAACGGCGTG	3532
Db	85592	GAAAGGAAAGATTAACCTTAGAGCCCAAGAGGTTGAGGCGGAGTATCATGAACATGCCACA	85533
Qy	3533	CTACACT-CAGTCTGGGTGAAGGTGCAGAGAGCTGTCTCAAAATTAATTAATTAATAAA	3591
Db	85532	CTGTACTCCAGCTGGGTGACAGAGCAACACTCTGTCTCAAAAAAAAAAAAAAAAAAAG	85473
Qy	3592	ATACCTTTTAAAAAACAATAATTAAATTTTAAAAACAACAACACTAGAGAT--GT	3649
Db	85472	AAATGAAAAAGAAAAAGAAAAAGAAAAAGAAAAACAACAAGAAACATGTGTAAA	85413
Qy	3650	TTGCAAATTTGATTAATTTGGAGGTCTATATCCGTGAGATTAATTAATTAATTTAGAGA	3709
Db	85412	AAGCAATTTGTTCCTGCGAATAATCACTATATCAAGTGTCTGTATCACTATACACA	85353
Qy	3710	GTTCTTCTCATTTCTTAGAGAGGTGAAATTTAAATATCAAGCTAGAAAGAACTAG	3769
Db	85352	AAACAAGCCTTCAGTTAAGGAAACTGGGTGAGGAGAGCAAGCAAGCAAGCAAGAGA	85293
Qy	3770	GGCTGCGCACTCCAAAGTGTGTCTCAAGACACAGCAGCATCAAGTAACTTGGGAAGTGT	3829
Db	85292	GGGTCTAAGGAAGAACTGAGAAATGTGTGACAGCAC-----TGTCTGGAGGCTGT	85240
Qy	3830	TAAATTTGAGAGTTTAGAGGCTTCAACCCAGACCTTAGAACA--GAATGTGATTAAAC	3887
Db	85239	CAGAAATGCAAGTCTTGGGCTTCTCCCGACCCACTGAAATCAGGGAGATATTTTAAAC	85180
Qy	3888	AAATTTCTAGGTGCTCAGAGGACAKTTAAACTTGAAGACTCTGCATCA	3939
Db	85179	AAACCTCAGAGTATCTTGCACAGGTTAAACTTGAAGACATCTGCACA	85128

Search completed: September 27, 2004, 14:32:31  
Job time : 22298 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: September 27, 2004, 07:46:34 | Search time 363 Seconds  
(without alignments)  
911.590 Million cell updates/sec

```
Title: US-10-030-294-2
Perfect score: 5960
Sequence: 1 cagctggtcaggagatgctg.....tttcgctcgatatctggatc 5960
```

Scoring table: IDENTITY\_NUC

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:\*

```

1: /cg2_6/prodata/2/ina/5A_COMB.seq: *
2: /cg2_6/prodata/2/ina/5B_COMB.seq: *
3: /cg2_6/prodata/2/ina/6A_COMB.seq: *
4: /cg2_6/prodata/2/ina/6B_COMB.seq: *
5: /cg2_6/prodata/2/ina/PTCTUS_COMB.seq:
6: /cg2_6/prodata/2/ina/backfillseq.seq:

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	187.6	3.1	8133	4	US-09-659-791A-10	Sequence 10, Appl
2	182.2	3.1	36651	4	US-09-738-094A-3	Sequence 3, Appl1
3	182.2	3.1	36651	4	US-09-964-469-3	Sequence 3, Appl1
4	176	3.0	36741	3	US-09-101-665-3	Sequence 3, Appl1
5	175.4	2.9	99500	4	US-09-798-096-10	Sequence 10, Appl
6	172.2	2.9	98844	4	US-09-791-111-10	Sequence 10, Appl
7	171.4	2.9	49312	4	US-09-671-117-485	Sequence 485, Appl
8	171.4	2.9	16692	4	US-09-818-512-3	Sequence 3, Appl1
9	171.2	2.9	21784	4	US-09-820-002-3	Sequence 3, Appl1
10	169.8	2.8	81001	4	US-09-750-580-1	Sequence 1, Appl1
11	169.4	2.8	40000	4	US-09-780-049-18	Sequence 18, Appl1
12	169	2.8	63000	4	US-09-780-172-18	Sequence 18, Appl1
13	169	2.8	11882	4	US-09-754-250-3	Sequence 3, Appl1
14	169	2.8	168575	4	US-09-426-290-1	Sequence 1, Appl1
15	168.6	2.8	55827	4	US-09-813-131A-3	Sequence 1, Appl1
16	167.6	2.8	246240	2	US-08-724-394A-20	Sequence 20, Appl
17	167.6	2.8	246240	2	US-08-724-394A-20	Sequence 20, Appl
18	167.6	2.8	246240	2	US-08-724-394A-22	Sequence 22, Appl
19	167	2.8	1332	4	US-09-584-568-1	Sequence 1, Appl1
20	166.8	2.8	74962	4	US-09-685-453A-3	Sequence 3, Appl1
21	166.6	2.8	1519	4	US-09-716-129-45	Sequence 45, Appl1
22	166.6	2.8	43950	4	US-09-735-34A-3	Sequence 3, Appl1
23	166.6	2.8	43950	4	US-10-060-332-3	Sequence 3, Appl1
24	166.6	2.8	161652	4	US-09-487-855A-40	Sequence 40, Appl
25	166.2	2.8	3882	4	US-09-548-797B-3	Sequence 3, Appl1
26	164.8	2.8	41584	4	US-09-536-059-1	Sequence 1, Appl1
27	164.2	2.8	29485	4	US-09-785-381-6	Sequence 6, Appl1

28	164	2.8	1001	4	US-09-671-317-170	Sequence 170, April
29	164	2.8	1001	4	US-09-671-317-450	Sequence 450, April
30	163.8	2.7	356	4	US-09-621-976-18402	Sequence 18402, A
31	163.8	2.7	940	4	US-09-659-791A-11	Sequence 11, April
32	163.8	2.7	1163	4	US-09-482-272-38	Sequence 38, April
33	163.8	2.7	84495	4	US-09-797-506-3	Sequence 3, April
34	163.2	2.7	405	4	US-09-621-976-9900	Sequence 9800, April
35	163.2	2.7	3271	4	US-09-548-797B-1	Sequence 1, April
36	163.2	2.7	392000	4	US-10-027-983-11	Sequence 11, April
37	162.8	2.7	35060	3	US-08-814-095-7	Sequence 7, April
38	162.8	2.7	75395	4	US-09-584-890-3	Sequence 3, April
39	162.4	2.7	720	3	US-09-318-448-5	Sequence 5, April
40	162.4	2.7	2907	4	US-09-023-655-1053	Sequence 1053, April
41	162.4	2.7	3390	4	US-09-448-797B-2	Sequence 2, April
42	161.8	2.7	20674	4	US-09-441-638-651	Sequence 651, April
43	161.8	2.7	46718	4	US-09-816-093-3	Sequence 3, April
44	161.6	2.7	63588	4	US-09-873-404-3	Sequence 3, April
45	161.6	2.7	162450	4	US-09-345-882-1	Sequence 1, April

## ALIGNMENTS

RESULT 1

; Sequence 10, Application US/09659791A  
Patent No. 6382808

### GENERAL INFORMATION

APPLICANT: Susan M. Fry

```

; FILE OF INVENTION: AMPLISENSE MODULATION OF CLOSTRIDIN EXPRESSION
;
; FILE REFERENCE: RTS-0156
;
; COMPANY AND LOCATION ATTACHED: TTC/06/CE0 7013

```

;; CURRENT APPLICATION NUMBER: 05/05/859,751  
;; CURRENT FILING DATE: 2000-09-11  
NUMBER OF SEQ ID NOS: 90

```

: NUMBER OF SEQ ID NOS: 3
: SEQ ID NO 10
: LENGTH: 8133

```

TYPE: DNA  
ORGANISM: Homo sapiens

US-09-659-791A-10

**Query Match**

Best Local Similarity 74.6%; Pred. NO. 3.5e-33;  
Matches 249; Conservative 0; Mismatches 84; Indels 1; Gaps 1

3264 TTATAATTTTTTAAAAAATAGATGAGTGTGCTCATGCGCTAATCCCAACATT 3322

Db 1057 TTTAAGAGGAAAAAACTGAGCTGGGCATGTTGCTCATGCTGTGATGCCAGCACTTT 111

3324 GGGAGCCGGGTCGGAGGATACCTTGAGTCCAGCACTTGAAGACCACTCAAGGCAAC 338

Db 1117 GAGAGGCCAAGCCGCGAGGATCATTTGAGGCCAGGAGTTTGAGACCAAGCTGGCCAGAT 1117

3384 AGCAAGACCCCATATCTTAAAAAACAAACAAAATACTCTGGTATGTTGTCT 3444

Db 117 AGTGAACCTGTCTTACAAATAAATACTTAAAAATTAGCCGGGTGTGGTGTCA 123

3444 CACCTTAGTCCAGCTACACAGAGCTGAGCCAGAGATCACTTGAGCCAGAGGT 3500

Db 1237 CACCTGACTCTCAGCTACTCCGAGAGCTCAGGCACAGAGTCACTGAGCCTGGAGTT 129

3504 TGAGGCTGCAGTGATCATGAACGGCTGTACT - CAGTTGGGTGACAGTCAAGAA 356

Db 1297 GGAGGCTGCGAGCTATGATTGCACCAATTGCATTCCAGCCTCGGCGMACGAGTGAAC 1356

Oy 3563 GCTGTCGAAMATATAAATTAATTAAAAATAC 3596

Db 1357 CCTGCTCTAATTAAAAATAATAAAAAATAC 1390

## RESULT 2

US-09-738-894A-3  
; Sequence 3, Application US/09738894A

Patent No. 6331423  
GENERAL INFORMATION: Karl et al  
APPLICANT: GUEGLER, ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEIN, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL000636  
CURRENT APPLICATION NUMBER: US/09/738,894A  
CURRENT FILING DATE: 2000-12-18  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO. 3  
LENGTH: 36651  
TYPE: DNA  
ORGANISM: Human  
FEATURES:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(36651)  
OTHER INFORMATION: n = A,T,C or G  
US-09-738-894A-3

Query Match	3.1%;	Score 182.2;	DB 4;	Length 36651;
Best Local Similarity	76.6%;	Pred. No. 1.1e-31;		
Matches 249;	Conservative 0;	Mismatches 73;	Indels 3;	Gaps 2.

Qy	3299	GTGTGTGTGGCTC	CATGCGCTGTAAT	ATCCAA	CAACTTGGGAA	CCGGAT	CCGGAGGTATGCT	3349
Db	11525	GTGCTGTACATCA	CGCCCTGTAA	TCCCAA	CACTTGGGAA	CCAAAGTGGG	AGATCCCTTT	11584
Qy	3350	GAGTCACACAGTT	TGAGACACAGT	CAGGCG	CAACACAG	CANAGACCC	CAATCTTAAAAAAAACA	3409
Db	11585	GAGGCGAAGATT	TGGAGCCAGCT	CTAGAGCA	ATAGCA	AGACTTTGTCTT	TTAAAAA--A	11642
Qy	3410	AAAACAAAACAAA	TTTACCTGGGTAT	TGGTGTGTCA	CCGTATGTC	CAAGCTAC	CACAGGAA	3469
Db	11643	AAAAAAAAAAAAA	TTTACTCTGGCAT	GTATGTATG	ATGACACCT	GTATATCCAG	CTACTCAGAGAG	11702
Qy	3470	GCTGAGGAGAAAG	ATCACTTATGAG	CCACAGAG	ATTGAGGCT	TCAGTGAT	CAATGAACGG	3529
Db	11703	GCTGAGGAGAGAG	ATTGCTTATGAG	CCCTGGAG	ATTGAGGCT	TCAGTGAT	CAATGAATCA	11762
Qy	3530	CTGTCTACACT	CACTGTGGGTGA	CAGTGC	CAGAGAGCT	GTCTCTCA	AAAAATTAATTAATTAATA	3588
Db	11763	CCACGTGCATC	CAGCTGTGA	TGACAGG	CAGAGACCT	GTCTCTTAA	CAAAACAAACAAACA	11822
Qy	3589	AAAACTACTTT	TAATAAAAA	CAAAAT	AT	3613		
Db	11823	AAAAACGAA	CAAAAA	CAAAATAT	11847			

```

1 RESULT 3
2 US-09-964-469-3
3 Sequence 3, Application US/09964469
4 Patent No. 6579709
5 GENERAL INFORMATION:
6 APPLICANT: GUEGLER, Karl et al
7 TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
8 TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
9 TITLE OF INVENTION: THREEOF
10 FILE REFERENCE: CLO000636D1V
11 CURRENT APPLICATION NUMBER: US/09/964,469
12 CURRENT FILING DATE: 2001-09-28
13 PRIOR APPLICATION NUMBER: 60/208,331
14 PRIOR FILING DATE: 2000-06-01
15 PRIOR APPLICATION NUMBER: 09/738,894
16 PRIOR FILING DATE: 2000-12-18
17 NUMBER OF SEQ ID NOS: 4
18 SOFTWARE: FastSeq for Windows Version 4.0
19 SEQ ID NO: 3
20 LENGTH: 36651
21 TYPE: DNA
22 ORGANISM: Human
23 FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: (1) ..(36651)
; OTHER INFORMATION: n = A,T,C or G
US-09-964-469-3

```

Query Match	3.1%	Score 182.2;	DB 4;	Length 3665;
Best Local Similarity	76.6%	Prod. No. 1.1e-31;		
Matches 249;	Conservative 0;	Mismatches 73;	Indels 3;	Gaps 2;

Oy	3290	GGTGGTGGCTCAGCCGTGATATCCCAACACTTTGGGAAAGCCGGATCCGGAGAGTAACTT	3349
Db	11525	GTGCTGTACTCAGGCGCTGATATCCCAACACTTTGGGAAAGCCAAAGTGTGGAGATCCCTT	11584
Oy	3350	GAGTCACAGAGTTTGAGACCAAGTCAGGGCAACAGACAGAGAGCCCATATCTTAAAAAAAACA	3409
Db	11585	GAGGCGAGAGTTTGAGACCAAGTCAGGGCAACAGAGAGAGCCCATATCTTAAAAA--A	1:6442
Oy	3410	AAACAAAACAAATTAATCTGGGATATGGTGTCTCAGCTGTGATGTCMAAGCTACACAGGAA	3469
Db	11643	AAAAAAAAAAAAATTAATCTGGGATATGGTGTCTCAGCTGTGATGTCMAAGCTACACAGAG	11702
Oy	3470	GGTAGGCGAAGAGATCACTTGAAGCCACAGAGGTGAAGCTCCAGTATCATCAAGAAAGCT	3529
Db	11703	GCTAGGCGAAGAGATCTTGAAGCCACAGAGGTGAAGCTCCAGTATCATCAAGAAAGCT	11762
Oy	3530	CTGCTTAATCTCAATCTCCGTATCAAGTCAGAGAGCTGTCTCAAAATTAATTAATTAATTA	3588
Db	11763	CACATGCTACTCAGCTGATGAGAGACAGGACAGACCTGTCTCTAAACAAACAAACAAACA	11822
Oy	3589	AAATTAATCTTTAAAAACAAATTAAT	3613
Db	11823	AAAAAAAAAAAAAAAAAAATTAAT	11847

RESULT 4  
 US-09-301-665-3  
 Sequence 3, Application US/09301665  
 Patent No. 6207876  
 GENERAL INFORMATION:  
 APPLICANT: KELLERMS.  
 APPLICANT: DATTA, RODNEY E.  
 APPLICANT: BLACKBURN, MICHAEL R.  
 TITLE OF INVENTION: ADENOSINE DEAMINASE DEFICIENT TRANSGENIC MICE AND  
 FILE REFERENCE: USH:243  
 CURRENT APPLICATION NUMBER: US/09/301,665  
 EARLIER FILING DATE: 1999-04-28  
 EARLIER APPLICATION NUMBER: 60/083,408  
 EARLIER FILING DATE: 1998-04-29  
 EARLIER APPLICATION NUMBER: 60/083,370  
 EARLIER FILING DATE: 1998-04-28  
 NUMBER OF SEQ ID NOS: 4  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO. 3  
 LENGTH: 36741  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-301-665-3

	Query Match	Similarity	3.0%;	Score 176;	DB 3;	length 36741;
	Best Local	Similarity	71.8%;	Pred. 2.9e-30;		
	Matches	244;	Conservative	0;	Mismatches	95; Indels 1; Gaps 1;
QY	3257	TAAACCTTTTAAATTTT	TTTTAAAAAATAAGATAGTGTGTCCTCAGCTTGATATCCA	3316		
DB	4879	TTACCTCTTCAAAATATAGTGTGTAAGTGCCCGCCAGAGGTGTGCTCAGCCTGTAAATCCA	4938			
QY	3317	AACATTTGGGAACCCCGGTCGGAGAGATCTGATCTGACATGATGTAAGACCAATGACAG	3376			
DB	4939	GCATTTGTGACGCGTGAAGCCGGAGATATTTGAGTCCAGAGTTTGAACATCAAGCTGG	4998			
QY	3377	GCACACAGAGAACCCCTATCTTAAAAAACAACAAAAACAATAATACCTGGATAGG	3438			

Thu Sep 30 11:51:28 2004

us-10-030-294-2.std.inl

**Page 3**

Db	4999	ACACAAAAC	TAGACC	CGCTCT	CCAAAAA	ATAAAAA	AGTTAG	GGAA	GTGTGTGG	5058
Oy	3437	TTGTGCTCA	CTGTAGT	CCAGCTAC	ACACAGGA	AGCTGAG	GCAGAA	AGATCA	TTTACGCC	3496
Db	5059	TGTGTGACT	CCGTAGT	CCAGCTAC	TACAGGA	GGCTGAGG	CGGAGAA	TGCTTTAG	CGCC	5118
Oy	3497	AGCAGGTG	AGAGCTG	CACTGAT	TCCATGA	AGCGCTGT	-AC	ACTCA	GTGTGGATAC	3555
Db	5119	AGCAGGTT	AGAGCTCA	GTGATG	GTGATG	TGCTCA	CTACCTT	CACCTT	CACCTTGG	5178
Oy	3556	GCACAGCT	GTTC	CAAAAT	TAATTA	TAAATTA	TA			3595
Db	5179	GCAGAGC	CTGTCT	CAAAAAA	AGAGAA	AGAAAGAA	A			5218

```

US-09-798-096-10/c
US-09-798-096-10
RESULT 5
US-09-798-096-10/c
Sequence 10. Application US/09798096
Patent No. 6399378
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
APPLICANT: Andrew T. Walt
TITLE OF INVENTION: AMTISENSE MODULATION OF RECOL2 EXPRESSION
FILE REFERENCE: RTS-0207
CURRENT APPLICATION NUMBER: US/09/798,096
CURRENT FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 69
SEQ ID NO 10
LENGTH: 99500
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
US-09-798-096-10

```

	Query Match	2.9%	Score 175.4	DB 4	Length 99500
	Best Local Similarity	73.2%	Pred. No. 5.9e-30		
	Matches 25;	Conservative 0	Mismatches 86	Indels 7	Gaps 2
OY	3276	AAAAAAAAATGATGAGTGTGTGGCTCATGCTGTAAATCCCAACACTTTGGGAAGCGCGGT	3335		
Dd	8073	AAAAAATTGACTGAGTAGTAAGAGGCTGCATGCTTGTAATCCCAAGCATTYYGGAGGCCAAGG	8014		
OY	3336	CGGAGGAGTAGCTTGAGTGCAGCACTTTGAGACCAAGTCAGGGCACACAGCAAGACCACA	3395		
Dd	8013	CAGAGTAGTACGCTTGAGCCCAAGGACTTTGAGACCAAGCTTAGGCGAACAAAGCAAGACCAG	7954		
OY	3396	TATCTTAAAATAAACCAAACCAAAACAAATTAATCTGGGATATGGTGTGCTCACTGATGATCC	3455		
Dd	7953	TATCTTCAAAAAATATA-----AAAGTTAGCCAAGATGTGTGTACCTCTGTATGCT	7900		
OY	3456	AAGCTTACACAGAGAGCTGAGCGAAGAATCACTTGAGCCCAAGACGTTGAGGCTGCAGT	3515		
Dd	7899	CACCTTCTCGGAGAGCTGAGCGAAGGATATCTGAGCCCAAGAGTTGAGAGCTGCAGT	7840		
OY	3516	GATCCATGAAACGGCTGCTACACT-CAGTCTGGGTGACAGTGCAGAGAACTGTCTCAAA	3574		
Dd	7839	GAGCAAGCATCTATTACTGCACTCCAGCTCGGATGACTGAGCAAAACCTGTCTGAAA	7780		
OY	3575	ATATATATAATTAATAATAATTAATTTTAAATAAAACAAATTAATTA	3621		
Dd	7779	ATATATATAATCATATCAATTAAGTTAATTAATAATAATAATAATAA	7733		

RESULT 6  
US-09-791-211-10/C  
; Sequence 10, Application US/09791211  
; Patent No. 6448080  
; GENERAL INFORMATION:  
; APPLICANT: Donna T. Ward  
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION  
; FILE REFERENCE: RTS-0205  
; CURRENT APPLICATION NUMBER: US/09/791.211

```

1 CURRENT FILING DATE: 2001-02-23
2 NUMBER OF SEQ ID NOS: 90
3 SEQ ID NO. 10
4 LENGTH: 98844
5 TYPE: DNA
6 ORGANISM: Homo sapiens
7 FEATURE:
8 NAME/KEY: unsure
9 LOCATION: 24962
10 OTHER INFORMATION: unknown
11 NAME/KEY: unsure
12 LOCATION: 64383
13 OTHER INFORMATION: unknown
14 NAME/KEY: unsure
15 LOCATION: 65468
16 OTHER INFORMATION: unknown
17 NAME/KEY: unsure
18 LOCATION: 65469
19 OTHER INFORMATION: unknown
20 NAME/KEY: unsure
21 LOCATION: 65470
22 OTHER INFORMATION: unknown
23 NAME/KEY: unsure
24 LOCATION: 65471
25 OTHER INFORMATION: unknown
26 NAME/KEY: unsure
27 LOCATION: 87130
28 OTHER INFORMATION: unknown
29 NAME/KEY: unsure
30 LOCATION: 89049
31 OTHER INFORMATION: unknown
32 OTHER INFORMATION:
33 US-09-791-211-10

```

[illegible]

RESULT 7  
US-09-671-317-485  
; Sequence 485, Application US/09671317  
; Patent No. 6528260  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bouguieleiret, Lydie  
; APPLICANT: Cohen, Amick  
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM



```
FILE REFERENCE: 62.US3.CIP
CURRENT APPLICATION NUMBER: US/09/671,317
CURRENT FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 09/536,178
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/IB00/00403
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/126,269
PRIOR FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: US 60/131,961
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 977
SOFTWARE: Patent.pm
SEQ ID NO 485
LENGTH: 49312
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5466..7466
OTHER INFORMATION: 5'regulatory region
NAME/KEY: exon
LOCATION: 7467..7725
OTHER INFORMATION: exon 1
NAME/KEY: exon
LOCATION: 20256..20355
OTHER INFORMATION: exon 2
NAME/KEY: exon
LOCATION: 36905..36975
OTHER INFORMATION: exon 3
NAME/KEY: exon
LOCATION: 45167..45248
OTHER INFORMATION: exon 4
NAME/KEY: exon
LOCATION: 45728..45965
OTHER INFORMATION: exon 5
NAME/KEY: misc_feature
LOCATION: 45966..49312
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 7564
OTHER INFORMATION: 10-286-289 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 7619
OTHER INFORMATION: 10-286-345 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 7649
OTHER INFORMATION: 10-286-375 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 17256
OTHER INFORMATION: 12-425-57 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 21590
OTHER INFORMATION: 12-421-135 : insertion of T
NAME/KEY: allele
LOCATION: 21595
OTHER INFORMATION: 12-421-140 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 36971
OTHER INFORMATION: 10-523-232 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 45214
OTHER INFORMATION: 10-289-201 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 45741
OTHER INFORMATION: 10-290-37 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 46028
OTHER INFORMATION: 10-290-326 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 46032
OTHER INFORMATION: 10-290-328 : deletion of G
NAME/KEY: primer_bind
LOCATION: 7276..7294
OTHER INFORMATION: 10-286.pu
NAME/KEY: primer_bind
LOCATION: 7676..7694
OTHER INFORMATION: 10-286.rp complement
NAME/KEY: primer_bind
LOCATION: 16839..16856
OTHER INFORMATION: 12-425.rp
NAME/KEY: primer_bind
LOCATION: 17297..17314
OTHER INFORMATION: 12-425.pu complement
NAME/KEY: primer_bind
LOCATION: 21456..21474
OTHER INFORMATION: 12-421.pu
NAME/KEY: primer_bind
LOCATION: 21886..21906
OTHER INFORMATION: 12-421.rp complement
NAME/KEY: primer_bind
LOCATION: 36740..36758
OTHER INFORMATION: 10-523.pu
NAME/KEY: primer_bind
LOCATION: 36997..37015
OTHER INFORMATION: 10-523.rp complement
NAME/KEY: primer_bind
LOCATION: 45020..45037
OTHER INFORMATION: 10-289.pu
NAME/KEY: primer_bind
LOCATION: 45413..45432
OTHER INFORMATION: 10-289.rp complement
NAME/KEY: primer_bind
LOCATION: 45705..45724
OTHER INFORMATION: 10-290.pu
NAME/KEY: primer_bind
LOCATION: 46104..46123
OTHER INFORMATION: 10-290.rp complement
NAME/KEY: primer_bind
LOCATION: 7545..7563
OTHER INFORMATION: 10-286-289.mis
NAME/KEY: primer_bind
LOCATION: 7565..7583
OTHER INFORMATION: 10-286-289.mis complement
NAME/KEY: primer_bind
LOCATION: 7600..7618
OTHER INFORMATION: 10-286-345.mis
NAME/KEY: primer_bind
LOCATION: 7620..7638
OTHER INFORMATION: 10-286-345.mis complement
NAME/KEY: primer_bind
LOCATION: 7630..7648
OTHER INFORMATION: 10-286-375.mis
NAME/KEY: primer_bind
LOCATION: 7650..7668
OTHER INFORMATION: 10-286-375.mis complement
NAME/KEY: primer_bind
LOCATION: 17239..17257
OTHER INFORMATION: 12-425-57.mis
NAME/KEY: primer_bind
LOCATION: 17259..17277
OTHER INFORMATION: 12-425-57.mis complement
NAME/KEY: primer_bind
LOCATION: 21576..21594
OTHER INFORMATION: 12-421-140.mis
NAME/KEY: primer_bind
LOCATION: 21596..21614
OTHER INFORMATION: 12-421-140.mis complement
NAME/KEY: primer_bind
LOCATION: 36952..36970
OTHER INFORMATION: 10-523-232.mis
NAME/KEY: primer_bind
LOCATION: 36972..36990
OTHER INFORMATION: 10-523-232.mis complement
NAME/KEY: primer_bind
LOCATION: 45195..45213
```



```

RESULT 10
US-09-750-580-1/c
; Sequence 1, Application US/09750580
; Patent No. 6455280
; GENERAL INFORMATION:
; APPLICANT: Yen, Frances
; APPLICANT: Denison, Blake
; APPLICANT: Bout, Barbara
; APPLICANT: Bihain, Bernard
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclet, Aymeric
; APPLICANT: Bougveleret, Lydie
; APPLICANT: Ebbeets-Reed, Dana
; APPLICANT: Salter-Cid, Luiza
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH
; FILE REFERENCE: 89 US2, CIP
; CURRENT APPLICATION NUMBER: US/09/750,580
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 09/559,362
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: PCT/IB00/0101
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: PCT/IB99/02058
; PRIOR FILING DATE: 1999-12-20
;

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	2.8%	Score 169.8	DB 4	Length 81001
	Best Local Similarity	67.7%	Pred. No. 1e-28		
	Matches 270	Conservative	0	Mismatches 122	Indels 7
	Gaps		2		
OY	3367	AAATTTTAAAAAATGACATCGTCGTCTCAGCTGTATTCACACATTGGG	3328		

```

RESULT 11
US-09-780-049-18/c
; Sequence 18, Application US/09780049
; Patent No. 6465250
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSOR MODULATION OF PROTEIN PHOSPHATASE 2 CATALYTIC SUBUNIT
; TITLE OF INVENTION: EXPRESSION
; FILE REFERENCE: RTS-0134
; CURRENT APPLICATION NUMBER: US/09/780,049
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 40000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-049-18

```

[illegible]









Query Match	3.1%	Score 187.6	DB 13	Length 8133
Best Local Similarity	74.6%	Pred. No. 3,4e-25		
Matches 249	Conservative 0	Mismatches 84	Indels 1	Gaps 1

  

Qy	3264	TTTTAAATTTTAAAAAAAATGAGTCTGTGCTGGCTGATGCTCTAAATGCCAACCTT	3323
Db	1057	TTTTAAAGAGAAATAAATCTGAGCTGGCAATGGTGGTCTCATGCTGTATGCAAGACTTT	11-6
Qy	3324	GGCAAGCCGGGTGTGGGAGATACCTTGAATCTCCAGACTTTGAGACAGTCAAGGCAACAC	3383
Db	1117	GAGAGGCCAAGGGGGAGAGATCATTTTGAAGCCAGAGTTTAAACACAGCTGGCCAAAGT	1176
Qy	3384	AGCAAGACCCCATATCTTAAAAAACAACAAACAAATTAAGCTGGTATGGTGTGCT	3443
Db	1177	AGTAGAGACCTGTCTCTCAAAAAATTTAAAACTTAAAGCCGGGTCTGCTGGTCA	1236
Qy	3444	CACCTGTAGTCCAAAGCTACAGAGNACTGAGCAGAGATCACTTGAAGCCAGAGGT	3503
Db	1237	CACCTGTAGTCTCAGCTAATCTGGGAGGCTCAGGCAAGAGATCACTGAGCTTGAGATT	1296
Qy	3504	TGAGGCTGACGTATCATGTAAACGGGTGGTCAACT-CAGTCTGGGTGAAGTGAAGAA	3562
Db	1297	GGAGGCTGACGTAGAGTATGATTCACACATTGATCTCAGGCTGGGCAACAGAGTGAGAC	1356
Qy	3563	GCTGTCCAAAAATATAATTAATTAATTAATTAATC	3596
Db	1357	CTGTCTTAATTTAAAAAATTAATTAATTAATTAATC	1390

  

RESULT 3
US-09-084-429-513
Sequence 513, Application US/09984429
Publication No. US20040010132A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 53 Human Secreted Proteins

```

Query Match 3.1%; Score 186; DB 11; Length 18861;
Best Local Similarity 76.1%; Pred. No. 1,2e-28;
Matches 242; Conservative 0; Mismatches 75; Indels 1; Gaps 1,

Qy 3265 TTAATTTTAAAAAATAGATGAGTGTGTGCTCATGCTGTATATCCCAACTTTG 3324
Db 5557 TAAAGTTGAAAAAATAATGACTAGGTGCAAGCGGCTCAATGTCTATATCCAGGCTTG 5616
Qy 3325 GGAAGCCGCGGTGCGAGAGATAGCTTGAGTCCAGCACTTTGAGACCAGTCAAGGCAACA 3384
Db 5617 GGAGGCCAGGCGGAGATATCACTTGAGCTCAGAGATTGTGAGACCAGCCTGAGCAACTA 5676
Qy 3385 GCAAGACCCCATATCTTAAATAAACAACAAATAATTACTGTGGTATGCTGTGCTC 3444
Db 5677 GCAGACTTCACTCTCTACTATAAAAAGAAAAGTTAAABATTTAGCCAGCGCTGTGCTGAC 5736
Qy 3445 ACCGTGAGTCCAGCTACACAGAGGTGAGGCAGAAAGATCACTTGAGCCAGAGGTT 3504
Db 5737 ACCCTAGTCCAGCTACTCAGAAAGCTGAGGACGAGATCACTTGAGCTGTGGAGGTT 5796
Qy 3505 GAGGCTGAGTATCCATGGAACGGCTGCTACACTCACTCTGGGTGACAGTCAAGAG 3563
Db 5797 GAGGCTGAGTATCCATGGAACGACACACTGCACTCACTCTGGGTGACAGTACAGCT 5856
Qy 3564 CTGCTCAAAAATAATAA 3581
Db 5857 CTGCTCAAAAAGCAAA 5874

RESULT 4
US-10-027-632-181221/c
: Sequence 181221, Application US/10027632
: Publication No. US20020198371A1
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OR INVENTION: Identification and Mapping of Single Nucleotide
: FILE REFERENCE: 108827.129
: CURRENT APPLICATION NUMBER: US/10/027,632
: PRIOR APPLICATION NUMBER: US 60/218,006
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20

```

```

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181221
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-181221

```

```

Query Match      3.1%; Score 185.6; DB 13; Length 611;
Best Local Similarity 71.6%; Pred. No. 2.1e-29;
Matches 272; Conservative 0; Mismatches 104; Indels 4; Gaps 2;

```

```

QY 3251 AGCTATTAACCTTTTAAATTTTAAATAATGATGAGTGTGCTGCTCATGCTGTA 3310
DB 410 AGCTTTCACCTGTTTAAAAAAGAAACAGGGGCTGACATGCTGCTCATGCTGTA 351
QY 3311 ATCCCAACCTTTGGGAGCGGCTGGGAGATAGCTTGAATCCAGCAGTTGACCA 3370
DB 350 ATCCAGACCTTTGGGAGCGGCTGGGAGATAGCTTGAATCCAGCAGTTGACCA 291
QY 3371 GTCAGGCAACACAGACAGACCCCATATCT---AAAAAACAACAACAATAATAC 3427
DB 290 GCTGAAACAATAGGAGACCTGCTCTACAAAAAATATATATATATATATATATAT 231
QY 3428 TGGGTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3487
DB 230 TGGGGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 171
QY 3488 CTGAGCCCAAGAGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3546
DB 170 CTGAGCCCTGGGAGGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 111
QY 3547 GGTGACAGTGAAGAGCTGTCTCAAAATTAATTAATTAATTAATTAATTAATTA 3606
DB 110 GGTGACAGTGAAGAGCTGTCTCAAAATTAATTAATTAATTAATTAATTAATTA 51
QY 3607 CAAAATTAATTAATTTTA 3626
DB 50 TAGATATGTTCCCTTTCTA 31

```

```

RESULT 5
US-10-027-632-181221/c
; Sequence 181221, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

```

```

; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181221
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-181221

```

```

Query Match      3.1%; Score 185.6; DB 16; Length 611;
Best Local Similarity 71.6%; Pred. No. 2.1e-29;
Matches 272; Conservative 0; Mismatches 104; Indels 4; Gaps 2;

```

```

QY 3251 AGCTATTAACCTTTTAAATTTTAAATAATGATGAGTGTGCTGCTCATGCTGTA 3310
DB 410 AGCTTTCACCTGTTTAAAAAAGAAACAGGGGCTGACATGCTGCTCATGCTGTA 351
QY 3311 ATCCCAACCTTTGGGAGCGGCTGGGAGATAGCTTGAATCCAGCAGTTGACCA 3370
DB 350 ATCCAGACCTTTGGGAGCGGCTGGGAGATAGCTTGAATCCAGCAGTTGACCA 291
QY 3371 GTCAGGCAACACAGACAGACCCCATATCT---AAAAAACAACAACAATAATAC 3427
DB 290 GCTGAAACAATAGGAGACCTGCTCTACAAAAAATATATATATATATATATATAT 231
QY 3428 TGGGTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3487
DB 230 TGGGGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 171
QY 3488 CTGAGCCCAAGAGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3546
DB 170 CTGAGCCCTGGGAGGCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 111
QY 3547 GGTGACAGTGAAGAGCTGTCTCAAAATTAATTAATTAATTAATTAATTAATTA 3606
DB 110 GGTGACAGTGAAGAGCTGTCTCAAAATTAATTAATTAATTAATTAATTAATTA 51
QY 3607 CAAAATTAATTAATTTTA 3626
DB 50 TAGATATGTTCCCTTTCTA 31

```

```

RESULT 6
US-10-027-632-115440/c
; Sequence 115440, Application US/10027632
; Publication No. US2002019837A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115440
; LENGTH: 3287

```







```

best LOCALSIMILARITY 76.63; FREQ: MO: 1.2e-27;
Matches 249; Conservative 0; Mismatches 73; Indels 3; Gaps 2;

```

Db 16899 TTGGAGGCTGAGCGGAGAATTGCTTGAAACGAGGAGCGAGCTTGCAAGAGAGCCGA 16840

QY 3523 GAAAGGCTGCTACACT-CAGCTGGGTGACAGTGCAGAGAGCTGCTCAAAATATATA 3581  
 DB 16839 GATCGCGCCACTGACACTGAGCTGGGTGACAGAGAGATCGCTCAAAATATATA 16780  
 QY 3582 ATAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3632  
 DB 16779 AAACAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 16729

## RESULT 14

US-10-681-199-8/c  
 ; Sequence 8, Application US/10681199  
 ; Publication No. US20040138441A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KERE, Julia  
 ; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLLEXIA  
 ; FILE REFERENCE: 0933-0214P  
 ; CURRENT APPLICATION NUMBER: US/10/681.199  
 ; CURRENT FILING DATE: 2003-10-09  
 ; NUMBER OF SEQ ID NOS: 42  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 50000  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: human DYX1C1 chromosomal gene region, nucleotides  
 ; US-10-681-199-8

Query Match 3.1%; Score 181.8; DB 17; Length 50000;  
 Best Local Similarity 68.4%; Pred. No. 1,76-27;  
 Matches 281; Conservative 0; Mismatches 127; Indels 3; Gaps 2;

QY 3223 CGTTAAATGATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3282  
 DB 17137 CTTCACAAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 17078  
 QY 3283 TAGATGAGTGTGCTGCTCATGCTGTAATCCCAACACTTTGGAGCCGGGTGGAGG 3342  
 DB 17077 AGCGTGGGCAATGCTGCTCATGCTGTAATCCCAACACTTTGGAGCCGGGTGGAGG 17018  
 QY 3343 ATACCTTGAAGTGCAGAGCTTGAACAGCTGAGGAGCAACAGAGAGCCCAATATCTAA 3402  
 DB 17017 ATCACTGAAGTGCAGAGCTTGAACAGCTGAGGAGCAACAGAGAGCCCAATATCTAA 16958  
 QY 3403 AAAACAAAACAAAACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3462  
 DB 16957 TAAACAAAACAAAACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 16900  
 QY 3463 ACAGGAGCTGAGGAGCAAGATCACTGAGCCAGAGGTTGAGGCTGAGGATGATCAT 3522  
 DB 16839 TTGGAGGCTGAGGAGGAGAAATGCTTGAACAGGAGGAGGAGGAGGAGGAGGAGGAG 16840  
 QY 3523 GAAGCGCTGTGTAAGT-CAGTGTGGGTGACAGTGCAGAGAGCTGCTCAAAATATATA 3581  
 DB 16839 GATCGCGCCACTGACACTGAGCTGGGTGACAGAGAGATCGCTCAAAATATATA 16780  
 QY 3582 ATAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3632  
 DB 16779 AAACAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 16729

## RESULT 15

US-09-901-152-3  
 ; Sequence 3, Application US/09901152  
 ; Publication No. US20030022624A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HU, Song et al.  
 ; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND  
 ; TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: CL001248  
 ; CURRENT APPLICATION NUMBER: US/09/901.152  
 ; CURRENT FILING DATE: 2001-07-10  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 58985  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)..(58985)  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; US-09-901-152-3

Query Match 3.0%; Score 181; DB 10; Length 58985;  
 Best Local Similarity 70.8%; Pred. No. 2,86-27;  
 Matches 271; Conservative 0; Mismatches 105; Indels 7; Gaps 2;

QY 3250 TAGCTATTAACCTTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3309  
 DB 8142 TGCGCTGAAACCTTTTGTCTTTTAAAGAAATGCGTGGGTGGGTGCTCAAGTCTGT 8201  
 QY 3310 AATCCAAACCTTTGGAGAGCGGGTGGGAGATAGCTTGAAGTCAAGATTGAGACC 3369  
 DB 8202 AATCCAAACCTTTGGAGAGCGGGTGGGAGATAGCTTGAAGTCAAGATTGAGACC 8261  
 QY 3370 AGTCAGGCAACAGCAGACAGCCATATCT-----AAAAACAAACAAACAAAT 3423  
 DB 8262 AGCTGCGCAACAGCAGACAGCCATATCT-----AAAAACAAACAAACAAAT 8321  
 QY 3424 TACCTGGATATGTTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATG 3483  
 DB 8322 TACCTGGATATGTTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATG 8381  
 QY 3484 ATCACTTGAAGTGCAGAGCTTGAACAGCTGAGGAGCAACAGAGAGCCCAATATCTAA 3542  
 DB 8382 ATCACTTGAAGTGCAGAGCTTGAACAGCTGAGGAGCAACAGAGAGCCCAATATCTAA 8441  
 QY 3543 TCTGGGTGACAGTGCAGAGAGCTGCTCAAAATTAATTAATTAATTAATTAATTAATTAAT 3602  
 DB 8442 CTGGGCGACAGCAGCAAGCTGCTCAAAATTAATTAATTAATTAATTAATTAATTAAT 8501  
 QY 3603 AAAACAAAATTAATTAATTTT 3625  
 DB 8502 GATGAGCAATGATGATGATTT 8524

Search completed: September 27, 2004, 19:08:45  
 Job time : 2481 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: September 27, 2004, 07:46:33 ; Search time 2018 seconds  
(without alignments)  
12423.582 Million cell updates/sec

Title: US-10-030-294-2

Perfect score: 5960  
Sequence: 1 cagctgttcggagatctcg.....tttggctgattcggatc 5960

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 212409941 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N\_Geneseq\_29Jan04:\*  
1: geneseq1980s:\*  
2: geneseq1990s:\*  
3: geneseq2000s:\*  
4: geneseq2001as:\*  
5: geneseq2001bs:\*  
6: geneseq2002s:\*  
7: geneseq2003as:\*  
8: geneseq2003bs:\*  
9: geneseq2003cs:\*  
10: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5960	100.0	5960	5	AAF25905 Human tum
2	5681.6	95.3	13940	5	AAF25915 Human tum
3	5676	95.2	5676	5	AAF25904 Human tum
4	5676	95.2	5676	5	AAF25907 Human tum
5	5676	95.2	5676	5	AAF25906 Human tum
6	193.4	3.2	118384	7	ABX56555 Human aut
7	189.4	3.2	1384	3	AACT74312 Human sec
8	187.6	3.1	8133	6	ABN99663 Human clu
9	185.4	3.1	47999	7	AAV252898 Human lwe
10	183	3.1	335199	9	ADC24703 Human wll
11	182.2	3.1	36651	8	AAV259728 Human dyc
12	181	3.0	58985	8	ABD16928 Human sec
13	181	3.0	143601	7	AAV252654 Human epi
14	181	3.0	26329	8	AAV252654 Human s
15	180	3.0	94770	8	AAV252654 Human s
16	180	3.0	94770	9	ADP72392 Human STR
17	180	3.0	94770	9	ADP72392 Human STR
18	179.8	3.0	108316	9	ADP72392 Human STR
19	179.6	3.0	65608	6	AB162910 Breast ca
20	179.6	3.0	65608	6	AB162910 Breast ca
21	179.6	3.0	65608	6	AB162910 Breast ca
22	178.5	3.0	186957	7	AB167668 Oesophagu
23	178.2	3.0	630	4	AA15941 Human bre

24	178.2	3.0	7726	5	ABA21079 Human ner
25	178.2	3.0	25423	4	AAK90279 Human dig
26	178.2	3.0	25423	4	AAI57656 Human col
27	178.2	3.0	25423	6	ABG99833 Genomic D
28	178.2	3.0	25423	6	ABG99833 Genomic D
29	178.2	3.0	25424	4	AAK90280 Human col
30	178.2	3.0	25424	4	AAI57657 Human col
31	178.2	3.0	25424	6	ABG99834 Human col
32	178.2	3.0	25424	6	ABG99837 Human col
33	177.8	3.0	16310	3	AAV21086 Human low
34	177.8	3.0	16310	3	AAV21086 Human low
35	177.8	3.0	16310	3	AAV21086 Human low
36	177.8	3.0	17634	3	AAV21087 Human low
37	177.8	3.0	17634	3	AAV21087 Human low
38	177.8	3.0	17634	7	ABZ96781 Human nuc
39	177.6	3.0	17588	9	ADP86824 Human nuc
40	177.2	3.0	9566	4	AAV25794 Human gen
41	177.2	3.0	9566	4	AAI37121 Human mus
42	177.2	3.0	9566	5	ABX74143 Human ner
43	177.2	3.0	9566	7	ABX74143 Human nov
44	177.2	3.0	9566	7	ABX60109 CDNA enco
45	177.2	3.0	183610	7	ACF62736 Cancer ba

## ALIGNMENTS

RESULT 1  
AAF25905  
ID AAF25905 standard; DNA: 5960 BP.  
AC AAF25905;  
DT 19-APR-2001 (first entry)  
XX Human tumor suppressor gene p51 promoter associated DNA SEQ ID 2.  
XX Tumor suppressor; p51; cell death; cell proliferation; cancer;  
XX cytoskeletal; gene therapy; screening; ds.  
XX Homo sapiens.  
OS WO200100818-A1.  
PN 04-JAN-2001.  
PD 28-JUN-2000; 2000WO-JP004261.  
PR 29-JUN-1999; 98JP-00183195.  
PX (NIPK) NIPPON KAYAKU KK.  
PY (SAKV) SAKAI T.  
PI Sakai T, Kagaya S, Sato T, Sukenaga Y, Fujii H;  
PI WPI, 2001-112452/12.  
PT New DNA sequences encoding the human p51 promoter domain for screening  
PT for compounds that modify p51 promoter activity and for diagnosing and  
PT treating cancer.  
PS Claim 1(4); Page 35-36; 60pp; Japanese.  
XX This invention describes novel DNA sequences (1) encoding the human p51  
XX promoter domain (which may include the 5'-untranslated sequence) or  
XX derived from it by addition, deletion and/or substitution of one or more  
XX bases. The invention also describes (1) expression plasmids including (1)  
XX (2) host cells transformed by (1); (3) DNA probes binding to all or  
XX part of (1); (4) cloning (1) using (3); (5) DNA (and RNA) corresponding to  
XX (1) inhibiting the expression of p51, which includes the antisense  
XX sequence of (1); (6) screening compounds for their ability to modify p51  
XX promoter activity, by observing their effect on cells transformed by (1);  
XX (7) compounds identified by (6); and (8) drug compositions containing

CC (7). The products of the invention have cytostatic activity and can be used for gene therapy. (1) is used to screen for compounds that modify p51 promoter activity. (1) and RNA corresponding to it are used to inhibit the expression of p51. (1) and the identified compounds are used for the diagnosis and treatment of cancer

Sequence 5960 BP, 1858 A, 1103 C, 1175 G, 1824 T, 0 U, 0 Other,

Query Match 100.0%; Score 5960; DB 5; Length 5960;

Beet Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 5960; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCTGTCAGGAGATGCTGGAAGAAAGCCCACTGCTTCTGCACTGAGGT 60  
 DB 1 CAGCTGTCAGGAGATGCTGGAAGAAAGCCCACTGCTTCTGCACTGAGGT 60

QY 61 GACTTGGAGGATGAGGTTGTCTGTTAAAGAACTGCACTCTTCTGCCCCAAT 120  
 DB 61 GACTTGGAGGATGAGGTTGTCTGTTAAAGAACTGCACTCTTCTGCCCCAAT 120

QY 121 TGGCTCTGTTCCCTTGCAAGCCCTCTTCTTGGGACCTCCCTTAAGGCACTCTTG 180  
 DB 121 TGGCTCTGTTCCCTTGCAAGCCCTCTTCTTGGGACCTCCCTTAAGGCACTCTTG 180

QY 121 TGGCTCTGTTCCCTTGCAAGCCCTCTTCTTGGGACCTCCCTTAAGGCACTCTTG 180  
 DB 121 TGGCTCTGTTCCCTTGCAAGCCCTCTTCTTGGGACCTCCCTTAAGGCACTCTTG 180

QY 181 ACATTAACCTAATTAATGTTTATTTGATGAATTTCACTGAGAGAGATGAG 240  
 DB 181 ACATTAACCTAATTAATGTTTATTTGATGAATTTCACTGAGAGAGATGAG 240

QY 241 GTCAATCAGAAAGACATGAGCTAAGGTTGCAAGTCTCTTTTCAATTAATTA 300  
 DB 241 GTCAATCAGAAAGACATGAGCTAAGGTTGCAAGTCTCTTTTCAATTAATTA 300

QY 301 AGCTATTCGAATACATTCAGTTTACTTAATTCAGGCCAGCTTACTCTTAAT 360  
 DB 301 AGCTATTCGAATACATTCAGTTTACTTAATTCAGGCCAGCTTACTCTTAAT 360

QY 361 GTCACTGAGCAATTAATTAAGTCCAAAGTTGGAAGATTAAGAGATCTCTCA 420  
 DB 361 GTCACTGAGCAATTAATTAAGTCCAAAGTTGGAAGATTAAGAGATCTCTCA 420

QY 421 CAGAACTTGGCTTCACTTTACTAAATTAAGATGTTGTTGTTGAGCTGACAT 480  
 DB 421 CAGAACTTGGCTTCACTTTACTAAATTAAGATGTTGTTGTTGAGCTGACAT 480

QY 481 CAGAACTTGGCTTCACTTTACTAAATTAAGATGTTGTTGTTGAGCTGACAT 480  
 DB 481 CAGAACTTGGCTTCACTTTACTAAATTAAGATGTTGTTGTTGAGCTGACAT 480

QY 541 TGGTGAATTTACTCACAACCTCTCTCTCTGAAAGTCTTCTCTCTCTCAATA 600  
 DB 541 TGGTGAATTTACTCACAACCTCTCTCTCTGAAAGTCTTCTCTCTCTCAATA 600

QY 601 AATGCAAGATGACCTGGAAGGCTAGACCTGAGGTTCAATTAACCTGACAAAG 660  
 DB 601 AATGCAAGATGACCTGGAAGGCTAGACCTGAGGTTCAATTAACCTGACAAAG 660

QY 661 AATGCAAGATGACCTGGAAGGCTAGACCTGAGGTTCAATTAACCTGACAAAG 660  
 DB 661 AATGCAAGATGACCTGGAAGGCTAGACCTGAGGTTCAATTAACCTGACAAAG 660

QY 720 TCACTTCTCTGATCTCATAGTCAAGGCTGCAAGCTCTTAAGGAATGCAATCAT 720  
 DB 720 TCACTTCTCTGATCTCATAGTCAAGGCTGCAAGCTCTTAAGGAATGCAATCAT 720

QY 721 CAGCTTAAAGCTCTCTGTTGAGCACTGTTTCTCACTGAGGCTGACATTAAT 780  
 DB 721 CAGCTTAAAGCTCTCTGTTGAGCACTGTTTCTCACTGAGGCTGACATTAAT 780

QY 781 TTTCTTGGTTAAAGAACTTTATTTTCTGAAAGCTCTCTCTCAATGATGAGG 840  
 DB 781 TTTCTTGGTTAAAGAACTTTATTTTCTGAAAGCTCTCTCTCAATGATGAGG 840

QY 841 CCAATCAGTCCAGATGACAGGCTCTCTGAGCAACATGCACTTTCTTTGGAA 900  
 DB 841 CCAATCAGTCCAGATGACAGGCTCTCTGAGCAACATGCACTTTCTTTGGAA 900

QY 901 TGAATCTTAAAGCTGAATAGCTGAAGTTCAAAAAGCTGTTGATTTGACTAG 960  
 DB 901 TGAATCTTAAAGCTGAATAGCTGAAGTTCAAAAAGCTGTTGATTTGACTAG 960

DB 901 TGAATCTTAAAGCTGAATAGCTGAAGTTCAAAAAGCTGTTGATTTGACTAG 960

QY 961 AGTGGCTTTGCAAGAGTACCTGCAATTCCTAATTTCTTAAGTCCGTAATTAAT 1020  
 DB 961 AGTGGCTTTGCAAGAGTACCTGCAATTCCTAATTTCTTAAGTCCGTAATTAAT 1020

QY 1021 CTTGGTACAGCCCTTCTGAGATGATGTTTCTTCCAACTGCTCTTAATAGCT 1080  
 DB 1021 CTTGGTACAGCCCTTCTGAGATGATGTTTCTTCCAACTGCTCTTAATAGCT 1080

QY 1081 GAATTTTCAATTTCTTTTCAATCAATTTTCAATTTTCTTCTGTTGTTGTT 1140  
 DB 1081 GAATTTTCAATTTCTTTTCAATCAATTTTCAATTTTCTTCTGTTGTTGTT 1140

QY 1141 GCTTTAGTGGCAGAACTGATTTCTGTTGTTTAAACCAAGGAATCTGATTA 1200  
 DB 1141 GCTTTAGTGGCAGAACTGATTTCTGTTGTTTAAACCAAGGAATCTGATTA 1200

QY 1201 TCTTCCCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260  
 DB 1201 TCTTCCCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260

QY 1261 TCTGAGATTAAGTATGATCAAGCTCTGATTAACCCAGGTTAGCCACTGTT 1320  
 DB 1261 TCTGAGATTAAGTATGATCAAGCTCTGATTAACCCAGGTTAGCCACTGTT 1320

QY 1321 AACATTTGCTCTCTCTGAGATTAAGAGTCTGGAAGAAAGATGATTAAT 1380  
 DB 1321 AACATTTGCTCTCTCTGAGATTAAGAGTCTGGAAGAAAGATGATTAAT 1380

QY 1381 CAAAGCTCAGAGAGCAAGCAAGATTCGAGATTTGATGATTAATCTTAA 1440  
 DB 1381 CAAAGCTCAGAGAGCAAGCAAGATTCGAGATTTGATGATTAATCTTAA 1440

QY 1441 AGTTTGTGAGACTTCTGAGCAATTAATTTGATTAAGCAAGATTTCTTAAT 1500  
 DB 1441 AGTTTGTGAGACTTCTGAGCAATTAATTTGATTAAGCAAGATTTCTTAAT 1500

QY 1501 AATCAACCAAGCAAGATTCAGAGCTCAGAGCTTAATTAAGATGATTAAT 1560  
 DB 1501 AATCAACCAAGCAAGATTCAGAGCTCAGAGCTTAATTAAGATGATTAAT 1560

QY 1561 AGTTGACATGATTAATTTCTGAGAAACATGATCAATGACAAATTAATCA 1620  
 DB 1561 AGTTGACATGATTAATTTCTGAGAAACATGATCAATGACAAATTAATCA 1620

QY 1621 AAGAGAAATAGTACCAATTAAGCTGCAATTTATGATTAATCTAGATTAAT 1680  
 DB 1621 AAGAGAAATAGTACCAATTAAGCTGCAATTTATGATTAATCTAGATTAAT 1680

QY 1681 CATAGTACCAAGTACCTGCAATGCTAATTAATTAATTAATTAATTAATTA 1740  
 DB 1681 CATAGTACCAAGTACCTGCAATGCTAATTAATTAATTAATTAATTAATTA 1740

QY 1741 AATAAATTTAGCTCTTCAATGAGCTTGAACCAATTAATTAATTAATTAAT 1800  
 DB 1741 AATAAATTTAGCTCTTCAATGAGCTTGAACCAATTAATTAATTAATTAAT 1800

QY 1801 GATGCAATTAATTAATTTCTTAAACAGAAAGTCTAATTTGGAAGAAATG 1860  
 DB 1801 GATGCAATTAATTAATTTCTTAAACAGAAAGTCTAATTTGGAAGAAATG 1860

QY 1861 AAAATATACATTAATTTCTTAAACAGAAAGTCTAATTTGGAAGAAATG 1920  
 DB 1861 AAAATATACATTAATTTCTTAAACAGAAAGTCTAATTTGGAAGAAATG 1920

QY 1921 TGAAGGAGATTCATTAAGTGAAGGATTTTGGGTAATTAAGGTTGTTG 1980  
 DB 1921 TGAAGGAGATTCATTAAGTGAAGGATTTTGGGTAATTAAGGTTGTTG 1980

QY 1981 ATCTTGAATGCAATGAAGCTACCTGATCTAATTTGATTAATTTGATTAAT 2040  
 DB 1981 ATCTTGAATGCAATGAAGCTACCTGATCTAATTTGATTAATTTGATTAAT 2040

Db 1981 AATCTGATTCGATGAAAGCTACGCTGCTTAAATGTCACCTCTCAGATGAAACCT 2040  
 Qy 2041 TCGAATTCGCAATTTCTATTGTAATGTAATTAATCTCTAAAGTAATCTTAAGAGTCA 2100  
 Db 2041 TCGAATTCGCAATTTCTATTGTAATGTAATTAATCTCTAAAGTAATCTTAAGAGTCA 2100  
 Qy 2101 AGTGTTCGCAATTTCTATTGTAATGTAATTAATCTCTAAAGTAATCTTAAGAGTCA 2160  
 Db 2101 AGTGTTCGCAATTTCTATTGTAATGTAATTAATCTCTAAAGTAATCTTAAGAGTCA 2160  
 Qy 2161 AATACATAACATGTAATTCATGCAATTCAGTTCGAATTTTCGTAAGTCTCTGAACAC 2220  
 Db 2161 AATACATAACATGTAATTCATGCAATTCAGTTCGAATTTTCGTAAGTCTCTGAACAC 2220  
 Qy 2221 AGAATGAAATTAAGATGCAAGCTGCTCAAGCTGTTCAATTCAGTACAGAGATG 2280  
 Db 2221 AGAATGAAATTAAGATGCAAGCTGCTCAAGCTGTTCAATTCAGTACAGAGATG 2280  
 Qy 2281 AGTCTATTCGCAATTTCTATTGTAATGTAATTAATCTCTAAAGTAATCTTAAGAGTCA 2340  
 Db 2281 AGTCTATTCGCAATTTCTATTGTAATGTAATTAATCTCTAAAGTAATCTTAAGAGTCA 2340  
 Qy 2341 GTCCAGATTAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2400  
 Db 2341 GTCCAGATTAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2400  
 Qy 2401 GAGGAGAAATTTTGTGCTTTGTAAGATGAGCAAGATGTAATGTAATGTAATGTAATGTAAT 2460  
 Db 2401 GAGGAGAAATTTTGTGCTTTGTAAGATGAGCAAGATGTAATGTAATGTAATGTAATGTAAT 2460  
 Qy 2461 TAAAGATTCCTGCTGAGAGGCAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2520  
 Db 2461 TAAAGATTCCTGCTGAGAGGCAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2520  
 Qy 2521 TGCACCTTAAGAGAAATGTCATGAAGGAGCAAGTGTAAATTAATTTTCATGAATGTA 2580  
 Db 2521 TGCACCTTAAGAGAAATGTCATGAAGGAGCAAGTGTAAATTAATTTTCATGAATGTA 2580  
 Qy 2581 GTGAGAGAAATTTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2640  
 Db 2581 GTGAGAGAAATTTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2640  
 Qy 2641 GAGATCAAAAGAAATGTCATGAAGGAGCAAGTGTAAATTAATTTTCATGAATGTA 2700  
 Db 2641 GAGATCAAAAGAAATGTCATGAAGGAGCAAGTGTAAATTAATTTTCATGAATGTA 2700  
 Qy 2701 AACACCTTCCTCTTAAGTACTTCTCTCCCTCCATCAAAATCTAAAGTCTTCA 2760  
 Db 2701 AACACCTTCCTCTTAAGTACTTCTCTCCCTCCATCAAAATCTAAAGTCTTCA 2760  
 Qy 2761 AACATCAATTAACAGGAGGTCATGTCAGAAAGGCAATTTGCTTTTCTGAAGTCTTAT 2820  
 Db 2761 AACATCAATTAACAGGAGGTCATGTCAGAAAGGCAATTTGCTTTTCTGAAGTCTTAT 2820  
 Qy 2821 GTACGTATTAATTAACAGGAGGTCATGTCAGAAAGGCAATTTGCTTTTCTGAAGTCTTAT 2880  
 Db 2821 GTACGTATTAATTAACAGGAGGTCATGTCAGAAAGGCAATTTGCTTTTCTGAAGTCTTAT 2880  
 Qy 2881 CTTCACATCAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940  
 Db 2881 CTTCACATCAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940  
 Qy 2941 AAGATGAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3000  
 Db 2941 AAGATGAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3000  
 Qy 3001 TTTCACTTGAATCGTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 3060  
 Db 3001 TTTCACTTGAATCGTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 3060  
 Qy 3061 CCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120  
 Db 3061 CCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120

Qy 3121 TTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180  
 Db 3121 TTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180  
 Qy 3181 ACTTTCAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3240  
 Db 3181 ACTTTCAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3240  
 Qy 3241 TAAATAGTTCAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300  
 Db 3241 TAAATAGTTCAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300  
 Qy 3301 CATGCTGTAATTCGCAATTTCTATTGTAATGTAATTAATGTAATGTAATGTAATGTAATGTAATGTAAT 3360  
 Db 3301 CATGCTGTAATTCGCAATTTCTATTGTAATGTAATTAATGTAATGTAATGTAATGTAATGTAATGTAAT 3360  
 Qy 3361 TTTGAGCCAGTCAAGGCAACGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3420  
 Db 3361 TTTGAGCCAGTCAAGGCAACGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3420  
 Qy 3421 AATTACCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480  
 Db 3421 AATTACCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480  
 Qy 3481 AGATCACTGTAAGCCGCAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3540  
 Db 3481 AGATCACTGTAAGCCGCAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3540  
 Qy 3541 AGTCTGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3600  
 Db 3541 AGTCTGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3600  
 Qy 3601 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660  
 Db 3601 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660  
 Qy 3661 TTAATTCGAGTCTATGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3720  
 Db 3661 TTAATTCGAGTCTATGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3720  
 Qy 3721 TTTCTGAGAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3780  
 Db 3721 TTTCTGAGAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3780  
 Qy 3781 CCAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3840  
 Db 3781 CCAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3840  
 Qy 3841 AGTCTGAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3900  
 Db 3841 AGTCTGAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3900  
 Qy 3901 GCTTCAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3960  
 Db 3901 GCTTCAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 3960  
 Qy 3961 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020  
 Db 3961 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020  
 Qy 4021 CCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4080  
 Db 4021 CCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4080  
 Qy 4081 GCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140  
 Db 4081 GCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140  
 Qy 4141 TCGAGCAAGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 4200  
 Db 4141 TCGAGCAAGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCT 4200



QY 4201 CACGATCTTAGCGCGCTCGGGCTAAGTAAAGAGAGATGTCAGTTTAAATAGCT 4260  
 DB 4201 CACGATCTTAGCGCGCTCGGGCTAAGTAAAGAGAGATGTCAGTTTAAATAGCT 4260  
 QY 4261 TCTCCCTTCATCTGCTGAGCAACAAATAAATATTTTATGAAACATTTTGAGT 4320  
 DB 4261 TCTCCCTTCATCTGCTGAGCAACAAATAAATATTTTATGAAACATTTTGAGT 4320  
 QY 4321 TAGATTACTTACAGGGAATGTCAATTTCTGAAAAGGCTTTGATGCTCAGAAC 4380  
 DB 4321 TAGATTACTTACAGGGAATGTCAATTTCTGAAAAGGCTTTGATGCTCAGAAC 4380  
 QY 4381 TTGAGCATCTACGATGTCACCTATTTACAGGTGCTCTGTGACTAGGAGGGA 4440  
 DB 4381 TTGAGCATCTACGATGTCACCTATTTACAGGTGCTCTGTGACTAGGAGGGA 4440  
 QY 4441 AGATGTGACACCAATGTAGTACCGTTAGATACAGAGTGTGTTTTTCCCTG 4500  
 DB 4441 AGATGTGACACCAATGTAGTACCGTTAGATACAGAGTGTGTTTTTCCCTG 4500  
 QY 4501 TTGAGTCTATCTCTAACTGAGCTCTGAAATCAATTTCAATTTCCAAATCCAGAA 4560  
 DB 4501 TTGAGTCTATCTCTAACTGAGCTCTGAAATCAATTTCAATTTCCAAATCCAGAA 4560  
 QY 4561 ACCGAGTAAAGTTTACAGCCCATATTCAAGAAAGAAATTAATTTTGTGTAGACT 4620  
 DB 4561 ACCGAGTAAAGTTTACAGCCCATATTCAAGAAAGAAATTAATTTTGTGTAGACT 4620  
 QY 4621 TCCGATATTTACACTGATTTGGAAATATATGACAAATTTATGGTTCCCTTCGAGTAG 4680  
 DB 4621 TCCGATATTTACACTGATTTGGAAATATATGACAAATTTATGGTTCCCTTCGAGTAG 4680  
 QY 4681 GTCAAGTCAGAGCAAAACCAAAACAGCAAAAACCTTACAGCATTAAGATAGCTGAG 4740  
 DB 4681 GTCAAGTCAGAGCAAAACCAAAACAGCAAAAACCTTACAGCATTAAGATAGCTGAG 4740  
 QY 4741 CGGCTGAGAGATTAATTAATTAATTAATTTTATTAACAGCAATTTGAAATTAAT 4800  
 DB 4741 CGGCTGAGAGATTAATTAATTAATTAATTTTATTAACAGCAATTTGAAATTAAT 4800  
 QY 4801 GTGCACTTCAGAAATTTCTACATTAATTAATTTCAATTTTAATCTTTAAGAAA 4860  
 DB 4801 GTGCACTTCAGAAATTTCTACATTAATTAATTTCAATTTTAATCTTTAAGAAA 4860  
 QY 4861 TTACTATTAATTAATTAATTAATTAATTAATTTTATTAATTAATTAATTAATTA 4920  
 DB 4861 TTACTATTAATTAATTAATTAATTAATTAATTTTATTAATTAATTAATTAATTA 4920  
 QY 4921 AGTTATTTTCTTTTATTCGGGTCAAGCAAGCTTCTAAGCAATGTGAAGCATATCT 4980  
 DB 4921 AGTTATTTTCTTTTATTCGGGTCAAGCAAGCTTCTAAGCAATGTGAAGCATATCT 4980  
 QY 4981 CTTTCTCTTAGTGAAGAGAGAGTGAAGTTCTAAGTTAAATTAATTAATTAATTAAT 5040  
 DB 4981 CTTTCTCTTAGTGAAGAGAGAGTGAAGTTCTAAGTTAAATTAATTAATTAATTAAT 5040  
 QY 5041 GTCTTCTCTTAGTGAAGAGAGAGTGAAGTTCTAAGTTAAATTAATTAATTAATTA 5100  
 DB 5041 GTCTTCTCTTAGTGAAGAGAGAGTGAAGTTCTAAGTTAAATTAATTAATTAATTA 5100  
 QY 5101 CGGGAG 5160  
 DB 5101 CGGGAG 5160  
 QY 5161 ACAAAACATTTTACGCGCGCAAAATAGTCAAGAAATCTCAAAATCAAAACAGATTCAG 5220  
 DB 5161 ACAAAACATTTTACGCGCGCAAAATAGTCAAGAAATCTCAAAATCAAAACAGATTCAG 5220  
 QY 5221 TACAG 5280  
 DB 5221 TACAG 5280  
 QY 5281 AAGTCCAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5340

DB 5281 AAGTCCAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5340  
 QY 5341 GAGATTCAG 5400  
 DB 5341 GAGATTCAG 5400  
 QY 5401 TACGTCAG 5460  
 DB 5401 TACGTCAG 5460  
 QY 5461 CCGTGAAG 5520  
 DB 5461 CCGTGAAG 5520  
 QY 5521 TCAAGAAAG 5580  
 DB 5521 TCAAGAAAG 5580  
 QY 5581 TGTCTGATAG 5640  
 DB 5581 TGTCTGATAG 5640  
 QY 5641 AGGATATTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5700  
 DB 5641 AGGATATTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5700  
 QY 5701 GGAATGATATTTGAAATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5760  
 DB 5701 GGAATGATATTTGAAATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5760  
 QY 5761 CGAGCGTGAAGTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5820  
 DB 5761 CGAGCGTGAAGTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5820  
 QY 5821 ATTATGATATTTGACAG 5880  
 DB 5821 ATTATGATATTTGACAG 5880  
 QY 5881 TATTATTTTGAAGTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5940  
 DB 5881 TATTATTTTGAAGTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5940  
 QY 5941 TTTTGTTGATTAATTTGATC 5960  
 DB 5941 TTTTGTTGATTAATTTGATC 5960  
 RESULT 2  
 AAF25915  
 ID AAF25915 standard; DNA: 13940 BP.  
 XX  
 AC AAF25915;  
 XX  
 XX 19-APR-2001 (first entry)  
 DT  
 XX Human tumor suppressor gene p51 promoter DNA fragment SEQ ID 12.  
 DE  
 XX Tumor suppressor; p51; cell death; cell proliferation; cancer;  
 KM cytostatic; gene therapy; screening; ds.  
 OS  
 XX Homo sapiens.  
 OS  
 PN W0200100818-A1.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 28-JUN-2000; 2000WC-JP004261.  
 XX  
 PR 29-JUN-1999; 99JP-00183195.  
 XX  
 PA (NIDK ) NIPON KAYAKU KK.  
 PA (SAKAI ) SAKAI T.

XX Sekai T, Kagaya S, Sato T, Sakenaga Y, Fujii H;  
XX  
XX WPI; 2001-112452/12.

XX New DNA sequences encoding the human p51 promoter domain for screening  
PT for compounds that modify p51 promoter activity and for diagnosing and  
PT treating cancer.

XX Example 1; Page 49-57; 60pp; Japanese.

XX This invention describes novel DNA sequences (1) encoding the human p51  
CC promoter domain (which may include the 5'-untranslated sequence) or  
CC derived from it by addition, deletion and/or substitution of one or more  
CC bases. The invention also describes (1) expression plasmids including (1)  
CC : (2) host cells transformed by (1); (3) DNA probes binding to a11 or  
CC part of (1); (4) cloning (1) using (3); (5) DNA (and RNA corresponding to  
CC 1c) inhibiting the expression of p51, which includes the antisense  
CC sequence of (1); (6) screening compounds for their ability to modify p51  
CC promoter activity, by observing their effect on cells transformed by (1);  
CC (7) compounds identified by (6); and (8) drug compositions containing  
CC (7). The products of the invention have cytotoxic activity and can be  
CC used for gene therapy. (1) is used to screen for compounds that modify  
CC p51 promoter activity. (1) and RNA corresponding to it are used to  
CC inhibit the expression of p51. (1) and the identified compounds are used  
CC for the diagnosis and treatment of cancer

XX Sequence 13940 BP; 3949 A; 2972 C; 3085 G; 3934 T; 0 U; 0 Other;

XX Query Match 95.3%; Score 5681.6; DB 5; Length 13940;

XX Best Local Similarity 99.3%; Pred. No. 0; Mismatches 39; Indels 0; Gaps 0;

XX Matches 5705; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 2 AGCTGTTGAGGATGCTGCAAGAAAGAACCCCACTTCTTGTGACACTGAGTGTG 61  
DB ATCTGTTGAGGATGCTGCAAGAAAGAACCCCACTTCTTGTGACACTGAGTGTG 110  
QY 62 ACTTGGAGGATGCTGCAAGAAAGAACCCCACTTCTTGTGACACTGAGTGTG 121  
DB ACTTGGAGGATGCTGCAAGAAAGAACCCCACTTCTTGTGACACTGAGTGTG 170  
QY 111 ACTTGGAGGATGCTGCAAGAAAGAACCCCACTTCTTGTGACACTGAGTGTG 170  
DB 111 ACTTGGAGGATGCTGCAAGAAAGAACCCCACTTCTTGTGACACTGAGTGTG 170  
QY 122 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181  
DB 122 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 230  
QY 171 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 230  
DB 171 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 230  
QY 182 CATTAACCTAATAATGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 241  
DB 182 CATTAACCTAATAATGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 290  
QY 231 CATTAACCTAATAATGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 290  
DB 231 CATTAACCTAATAATGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 290  
QY 242 TCMAATGCAAGAGCAAGGCTAAGTGCATGACCTGCTTTCATGTAATTAA 301  
DB 242 TCMAATGCAAGAGCAAGGCTAAGTGCATGACCTGCTTTCATGTAATTAA 301  
QY 291 TCMAATGCAAGAGCAAGGCTAAGTGCATGACCTGCTTTCATGTAATTAA 350  
DB 291 TCMAATGCAAGAGCAAGGCTAAGTGCATGACCTGCTTTCATGTAATTAA 350  
QY 302 GTCATTCGATTAACCTATGTTACTTAAAGTTTGAAGCCGCTTAACTCTAATGATG 361  
DB 302 GTCATTCGATTAACCTATGTTACTTAAAGTTTGAAGCCGCTTAACTCTAATGATG 410  
QY 351 GTCATTCGATTAACCTATGTTACTTAAAGTTTGAAGCCGCTTAACTCTAATGATG 410  
DB 351 GTCATTCGATTAACCTATGTTACTTAAAGTTTGAAGCCGCTTAACTCTAATGATG 410  
QY 362 TCAGACTGTAGCAAAATATTTAGTCCAAAGTTGCAAGAGTATGCTCTCAATGAC 421  
DB 362 TCAGACTGTAGCAAAATATTTAGTCCAAAGTTGCAAGAGTATGCTCTCAATGAC 470  
QY 411 TCAGACTGTAGCAAAATATTTAGTCCAAAGTTGCAAGAGTATGCTCTCAATGAC 470  
DB 411 TCAGACTGTAGCAAAATATTTAGTCCAAAGTTGCAAGAGTATGCTCTCAATGAC 470  
QY 422 AGAAGCTTGGCTTGCACCTTACTTAAATATGAGATTTGTGTTAGTTCGACCTATGAC 481  
DB 422 AGAAGCTTGGCTTGCACCTTACTTAAATATGAGATTTGTGTTAGTTCGACCTATGAC 530  
QY 471 AGAAGCTTGGCTTGCACCTTACTTAAATATGAGATTTGTGTTAGTTCGACCTATGAC 530  
DB 471 AGAAGCTTGGCTTGCACCTTACTTAAATATGAGATTTGTGTTAGTTCGACCTATGAC 530  
QY 482 AGAAGCTTGGCTTGCACCTTACTTAAATATGAGATTTGTGTTAGTTCGACCTATGAC 541  
DB 482 AGAAGCTTGGCTTGCACCTTACTTAAATATGAGATTTGTGTTAGTTCGACCTATGAC 590  
QY 531 AGAAGCTTGGCTTGCACCTTACTTAAATATGAGATTTGTGTTAGTTCGACCTATGAC 590  
DB 531 AGAAGCTTGGCTTGCACCTTACTTAAATATGAGATTTGTGTTAGTTCGACCTATGAC 590  
QY 542 GGTGAATTTACTTCCACCACTCTCTCTCTTGAAGGTTCTTCTCTCTCTCTCTCTCTA 601  
DB 542 GGTGAATTTACTTCCACCACTCTCTCTCTTGAAGGTTCTTCTCTCTCTCTCTCTCTA 650  
QY 591 GGTGAATTTACTTCCACCACTCTCTCTTGAAGGTTCTTCTCTCTCTCTCTCTCTA 650  
DB 591 GGTGAATTTACTTCCACCACTCTCTTGAAGGTTCTTCTCTCTCTCTCTCTCTA 650

QY 602 ATCAGATGACCTGCAAGGCTAGAGACCTGAGTTCAGTTACCTGACACAAAGAAAT 661  
DB 602 ATCAGATGACCTGCAAGGCTAGAGACCTGAGTTCAGTTACCTGACACAAAGAAAT 710  
QY 651 ATCAGATGACCTGCAAGGCTAGAGACCTGAGTTCAGTTACCTGACACAAAGAAAT 710  
DB 651 ATCAGATGACCTGCAAGGCTAGAGACCTGAGTTCAGTTACCTGACACAAAGAAAT 710  
QY 662 CAGTTCTCTGATTCATATGATCAAGAGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 721  
DB 662 CAGTTCTCTGATTCATATGATCAAGAGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 770  
QY 711 CAGTTCTCTGATTCATATGATCAAGAGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 770  
DB 711 CAGTTCTCTGATTCATATGATCAAGAGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 770  
QY 722 TGCCTTAAGCTCTTGTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781  
DB 722 TGCCTTAAGCTCTTGTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 830  
QY 771 TGCCTTAAGCTCTTGTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 830  
DB 771 TGCCTTAAGCTCTTGTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 830  
QY 782 TCCCTGCTTGAAGCAAGCTTATTTCTTGAAGAACTCTGCTGCTGCTGCTGCTGCTG 841  
DB 782 TCCCTGCTTGAAGCAAGCTTATTTCTTGAAGAACTCTGCTGCTGCTGCTGCTGCTG 890  
QY 831 TCCCTGCTTGAAGCAAGCTTATTTCTTGAAGAACTCTGCTGCTGCTGCTGCTGCTG 890  
DB 831 TCCCTGCTTGAAGCAAGCTTATTTCTTGAAGAACTCTGCTGCTGCTGCTGCTGCTG 890  
QY 842 CATCAGTCAATGATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901  
DB 842 CATCAGTCAATGATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950  
QY 891 CATCAGTCAATGATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950  
DB 891 CATCAGTCAATGATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950  
QY 902 GAATCTTAAGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 961  
DB 902 GAATCTTAAGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1010  
QY 951 GAATCTTAAGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1010  
DB 951 GAATCTTAAGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1010  
QY 962 GTGCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1021  
DB 962 GTGCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1070  
QY 1011 GTGCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1070  
DB 1011 GTGCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1070  
QY 1022 CTGCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1081  
DB 1022 CTGCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1130  
QY 1071 CTGCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1130  
DB 1071 CTGCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1130  
QY 1082 AATTTGAATTTCTTTCATATGATCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1141  
DB 1082 AATTTGAATTTCTTTCATATGATCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1190  
QY 1131 AATTTGAATTTCTTTCATATGATCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1190  
DB 1131 AATTTGAATTTCTTTCATATGATCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1190  
QY 1142 GCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1201  
DB 1142 GCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1250  
QY 1191 GCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1250  
DB 1191 GCTTGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1250  
QY 1202 CTTTCCCTTTTAAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA 1261  
DB 1202 CTTTCCCTTTTAAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA 1310  
QY 1251 CTTTCCCTTTTAAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA 1310  
DB 1251 CTTTCCCTTTTAAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA 1310  
QY 1262 CTGGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1321  
DB 1262 CTGGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1370  
QY 1311 CTGGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1370  
DB 1311 CTGGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1370  
QY 1322 AATTTGCTCTCTCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1381  
DB 1322 AATTTGCTCTCTCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1430  
QY 1371 AATTTGCTCTCTCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1430  
DB 1371 AATTTGCTCTCTCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1430  
QY 1382 AAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1441  
DB 1382 AAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1490  
QY 1431 AAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1490  
DB 1431 AAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1490  
QY 1442 GTTGTGGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
DB 1442 GTTGTGGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1550  
QY 1491 GTTGTGGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1550  
DB 1491 GTTGTGGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1550  
QY 1502 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1561  
DB 1502 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1610  
QY 1551 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1610  
DB 1551 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1670  
QY 1562 GTTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1621  
DB 1562 GTTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1670  
QY 1611 GTTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1670  
DB 1611 GTTGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1670  
QY 1622 CAGGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1681  
DB 1622 CAGGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1730  
QY 1671 CAGGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1730  
DB 1671 CAGGAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1730  
QY 1682 ATAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1741  
DB 1682 ATAAGGCTGATGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCA 1741





XX 29-JUN-1999; 99JP-00183195.  
 XX (NIPK) NIPKON KAYAKU KK.  
 PA (SAKAI) SAKAI T.  
 XX  
 PI Sakai T, Kagaya S, Sato T, Bokenaga Y, Fujii H,  
 XX WPI, 2001-112452/12.  
 XX

XX New DNA sequences encoding the human p51 promoter domain for screening  
 PT for compounds that modify p51 promoter activity and for diagnosing and  
 PT treating cancer.  
 XX

XX Claim 1(1); Page 31-34; 60pp; Japanese.  
 XX

CC This invention describes novel DNA sequences (1) encoding the human p51  
 CC promoter domain (which may include the 5'-untranslated sequence) or  
 CC derived from it by addition, deletion and/or substitution of one or more  
 CC bases. The invention also describes (1) expression plasmids including (1)  
 CC ; (2) host cells transformed by (1); (3) DNA probes binding to all or  
 CC part of (1); (4) cloning (1) using (3); (5) DNA (and RNA corresponding to  
 CC it) inhibiting the expression of p51, which includes the antisense  
 CC sequence of (1); (6) screening compounds for their ability to modify p51  
 CC promoter activity, by observing their effect on cells transformed by (1);  
 CC (7) compounds identified by (6); and (8) drug compositions containing  
 CC used for gene therapy. (1) is used to screen for compounds that can be  
 CC p51 promoter activity. (1) and RNA corresponding to it are used to  
 CC inhibit the expression of p51. (1) and the identified compounds are used  
 CC for the diagnosis and treatment of cancer  
 XX

XX Sequence 5676 BF: 1784 A; 1053 C; 1116 G; 1723 T; 0 U; 0 Other;

Query Match 95.2%; Score 5676; DB 5; Length 5676;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCTGTTCAAGGATGCTGGAAGAAAGAGCCCAACCACTGCTTGTGACACTGGGTGT 60  
 DB 1 CAGCTGTTCAAGGATGCTGGAAGAAAGAGCCCAACCACTGCTTGTGACACTGGGTGT 60  
 QY 61 GACTTGGAGGGATATCAGGTTGTCTGTTAAAGAAAGTGCACACTCTTCTGCCCCAT 120  
 DB 61 GACTTGGAGGGATATCAGGTTGTCTGTTAAAGAAAGTGCACACTCTTCTGCCCCAT 120  
 QY 121 TGGCTCTGTTCCCTTGCATGCGCTCTTCCCTGGGACACTCCCTTAAGCATCTTTTG 180  
 DB 121 TGGCTCTGTTCCCTTGCATGCGCTCTTCCCTGGGACACTCCCTTAAGCATCTTTTG 180  
 QY 121 TGGCTCTGTTCCCTTGCATGCGCTCTTCCCTGGGACACTCCCTTAAGCATCTTTTG 180  
 DB 121 TGGCTCTGTTCCCTTGCATGCGCTCTTCCCTGGGACACTCCCTTAAGCATCTTTTG 180  
 QY 181 ACATTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
 DB 181 ACATTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
 QY 241 GTCAATACAGAAAGACATGAGTAAAGTGAATGCACTGCTTTTCAATTAATTAATTA 300  
 DB 241 GTCAATACAGAAAGACATGAGTAAAGTGAATGCACTGCTTTTCAATTAATTAATTA 300  
 QY 301 AGTGAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360  
 DB 301 AGTGAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360  
 QY 361 GTCAAGCTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420  
 DB 361 GTCAAGCTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420  
 QY 421 CAGAACTTGGCTTCCCTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480  
 DB 421 CAGAACTTGGCTTCCCTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480  
 QY 481 CAGAACTTGGCTTCCCTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540  
 DB 481 CAGAACTTGGCTTCCCTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540

QY 541 TGGTAAATTTATCTGACACACCTCTCTCTTGAAGGTTCTTCTGCTCTCTCACTATA 600  
 DB 541 TGGTAAATTTATCTGACACACCTCTCTCTTGAAGGTTCTTCTGCTCTCTCACTATA 600  
 QY 601 AATGCAAGATGACCTGGAAGGCTAGAGACTGAGGTTGAGTACCTGACAAAGGAT 660  
 DB 601 AATGCAAGATGACCTGGAAGGCTAGAGACTGAGGTTGAGTACCTGACAAAGGAT 660  
 QY 661 TCAATTTCTGATCTCATATGATGACAGAGTGCACAGCTCTACAGAACTGCAAGATCAT 720  
 DB 661 TCAATTTCTGATCTCATATGATGACAGAGTGCACAGCTCTACAGAACTGCAAGATCAT 720  
 QY 721 CTGCTTAAAGCTCTTGTGAGGCAATCTGTTTTCATCTGCTGACCTATTTGCTCT 780  
 DB 721 CTGCTTAAAGCTCTTGTGAGGCAATCTGTTTTCATCTGCTGACCTATTTGCTCT 780  
 QY 781 TTCTTGGTTTACAGAACTTTTATTTTCTTGAAGAACTCTGCTGACATGATGAGG 840  
 DB 781 TTCTTGGTTTACAGAACTTTTATTTTCTTGAAGAACTCTGCTGACATGATGAGG 840  
 QY 841 CCAATGACCAATGATGACAGGCTCTCTGCGCAAACTGAGATCTTTTGGGAAAT 900  
 DB 841 CCAATGACCAATGATGACAGGCTCTCTGCGCAAACTGAGATCTTTTGGGAAAT 900  
 QY 901 TGAATCTTAAGCTGAATGCTGAAAGTTCAAAAGGCTGTAATCTGATACGCTTAC 960  
 DB 901 TGAATCTTAAGCTGAATGCTGAAAGTTCAAAAGGCTGTAATCTGATACGCTTAC 960  
 QY 961 AGTGGCTTGCAGAAAGTACGTCATGCTATTTCTTAAGCCTGAAATTAATTAAT 1020  
 DB 961 AGTGGCTTGCAGAAAGTACGTCATGCTATTTCTTAAGCCTGAAATTAATTAAT 1020  
 QY 1021 CCGGTTACAGCCCTTTCTGAGATGTTGTTTTCCTCAACTGCTCTTAATAGTCTGT 1080  
 DB 1021 CCGGTTACAGCCCTTTCTGAGATGTTGTTTTCCTCAACTGCTCTTAATAGTCTGT 1080  
 QY 1081 GAATTTTCATATTTCTTTTCAATTTTCATTTTCATTTTGTGTTGTTGTTGTTT 1140  
 DB 1081 GAATTTTCATATTTCTTTTCAATTTTCATTTTCATTTTGTGTTGTTGTTGTTT 1140  
 QY 1141 GGGTTTGGTGAAGGAGAACTGCTCTGTTTGAATTAATCCAGAAATCCCTGATTAATCA 1200  
 DB 1141 GGGTTTGGTGAAGGAGAACTGCTCTGTTTGAATTAATCCAGAAATCCCTGATTAATCA 1200  
 QY 1201 TCCCTTCCCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260  
 DB 1201 TCCCTTCCCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260  
 QY 1261 TCTGGAGTAAAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320  
 DB 1261 TCTGGAGTAAAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320  
 QY 1321 AAATTTTCTCTCTCTCTGAGTTTAAAGCTCTTGGAGAAAGAAAGTAAATTTTTC 1380  
 DB 1321 AAATTTTCTCTCTCTCTGAGTTTAAAGCTCTTGGAGAAAGAAAGTAAATTTTTC 1380  
 QY 1381 CAAAGCTGAGAGAGACCAAGAACTGAGATTAATGATGATTAATCTTAATTAAT 1440  
 DB 1381 CAAAGCTGAGAGAGACCAAGAACTGAGATTAATGATGATTAATCTTAATTAAT 1440  
 QY 1441 AGTTTGTGAGCTTCTGAGCAATTAATTTGTTTCTTAAGCAAGTTTCAATTTAATAGCT 1500  
 DB 1441 AGTTTGTGAGCTTCTGAGCAATTAATTTGTTTCTTAAGCAAGTTTCAATTTAATAGCT 1500  
 QY 1501 AAACAAACAAACAAAGATTCACAGGTTTCAGAGTTAATTAAGATGAAATTAATCA 1560  
 DB 1501 AAACAAACAAACAAAGATTCACAGGTTTCAGAGTTAATTAAGATGAAATTAATCA 1560  
 QY 1561 AGTTGACACATGATATATCTCAAGAACTGACATCAATGACAAAGAAATCAAGAC 1620  
 DB 1561 AGTTGACACATGATATATCTCAAGAACTGACATCAATGACAAAGAAATCAAGAC 1620







Db 3781 CCAAGTGTGTCCAGACACAGACATCAAGTACCTGGAAAGTGTAGAAATGACG 3840  
 Qy 3841 AGTTTATGGCTTCAACCCAGACCTTCTGACCAAGATCTGTATTAACAGATTTCTAGT 3900  
 Db 3841 AGTTTATGGCTTCAACCCAGACCTTCTGACCAAGATCTGTATTAACAGATTTCTAGT 3900  
 Qy 3901 GCTTCAGGGCAGATTAAACTTGGAGAGCTCTGACAGAAATCTCACTCCACCTTC 3960  
 Db 3901 GCTTCAGGGCAGATTAAACTTGGAGAGCTCTGACAGAAATCTCACTCCACCTTC 3960  
 Qy 3961 ATTATTAATGSAATCTTGGGCTGTGTGACAGAGAAATTGATTAATTTTAAATTCAGAA 4020  
 Db 3961 ATTATTAATGSAATCTTGGGCTGTGTGACAGAGAAATTGATTAATTTTAAATTCAGAA 4020  
 Qy 4021 CCTCTATTTAGTCACTATATTTGCTATATGCAAGGAGAAAGCCAAATCTTTTAACT 4080  
 Db 4021 CCTCTATTTAGTCACTATATTTGCTATATGCAAGGAGAAAGCCAAATCTTTTAACT 4080  
 Qy 4081 GCAATTAACAAATCTATTAATTAATTAAGTAAATCTTCCCTTAAGTTTACATTTTG 4140  
 Db 4081 GCAATTAACAAATCTATTAATTAATTAAGTAAATCTTCCCTTAAGTTTACATTTTG 4140  
 Qy 4141 TGGAGCAGAGCTGTATATTTGGCTGGGCTGAGGCGGCTGTGTGATTTTCAATTT 4200  
 Db 4141 TGGAGCAGAGCTGTATATTTGGCTGGGCTGAGGCGGCTGTGTGATTTTCAATTT 4200  
 Qy 4201 CACGATGTTAGCGCTCTGGGCTTAAAGGAGAGAAATGCAAGTTTAAATAGCT 4260  
 Db 4201 CACGATGTTAGCGCTCTGGGCTTAAAGGAGAGAAATGCAAGTTTAAATAGCT 4260  
 Qy 4261 TCTCCCTTCATCTGGCTGAGCAACAAATTAATTTTAAATTAAGAAACATTTTGGT 4320  
 Db 4261 TCTCCCTTCATCTGGCTGAGCAACAAATTAATTTTAAATTAAGAAACATTTTGGT 4320  
 Qy 4321 TAGATTAATTAAGGAAATGCAAAATTTCTGAAAGGCTTTAGATTTCTCAAC 4380  
 Db 4321 TAGATTAATTAAGGAAATGCAAAATTTCTGAAAGGCTTTAGATTTCTCAAC 4380  
 Qy 4381 TTTGACATCTAGTATGATCTATTTTCAAGTGTCTCTGATCTAGAGGCTGAAAGGA 4440  
 Db 4381 TTTGACATCTAGTATGATCTATTTTCAAGTGTCTCTGATCTAGAGGCTGAAAGGA 4440  
 Qy 4441 AGATGTAATCTACACAGATTTAGGACCGTATGATGACAGAGGCTTTTTCCTCCG 4500  
 Db 4441 AGATGTAATCTACACAGATTTAGGACCGTATGATGACAGAGGCTTTTTCCTCCG 4500  
 Qy 4501 TTTGAGCTATCTTAACTGAGCTTCTGATGATGATTTCAATTTCCAAATCCAGAA 4560  
 Db 4501 TTTGAGCTATCTTAACTGAGCTTCTGATGATGATTTCAATTTCCAAATCCAGAA 4560  
 Qy 4561 ACCAGAGTAAAGTTTACAGCCCATATTCAGAAAGAAATTAATTTTGTGTAGACT 4620  
 Db 4561 ACCAGAGTAAAGTTTACAGCCCATATTCAGAAAGAAATTAATTTTGTGTAGACT 4620  
 Qy 4621 TCCGATTAATTAAGCATTGAGGAAATATATTAAGATTTTCCCTTCCGAAATG 4680  
 Db 4621 TCCGATTAATTAAGCATTGAGGAAATATATTAAGATTTTCCCTTCCGAAATG 4680  
 Qy 4681 GTCAAGCTCAAGCAAAACCAAAACAGCAAAATCTGTAAGACATTAAGATAGTGAG 4740  
 Db 4681 GTCAAGCTCAAGCAAAACCAAAACAGCAAAATCTGTAAGACATTAAGATAGTGAG 4740  
 Qy 4741 CCGACTGAGAGATTAATACTAGATATTTTATTAACAGCAATTTGAAATTAATTT 4800  
 Db 4741 CCGACTGAGAGATTAATACTAGATATTTTATTAACAGCAATTTGAAATTAATTT 4800  
 Qy 4801 GTGAGCTGAGATATTTGAGCAATTAATTAATTTTCAATTTTAAATCTTTAAGAAA 4860  
 Db 4801 GTGAGCTGAGATATTTGAGCAATTAATTAATTTTCAATTTTAAATCTTTAAGAAA 4860  
 Qy 4861 TTACTATATTAATTAATTAATTAATTAATTTTGAAGTATTTTAACTCAATTA 4920  
 Db 4861 TTACTATATTAATTAATTAATTAATTTTGAAGTATTTTAACTCAATTA 4920

Qy 4921 AGTTATTTTCTTTATTTGGGCTCAGGCAAACTTCTTAAGGAGATGTAAGAGGATATCT 4980  
 Db 4921 AGTTATTTTCTTTATTTGGGCTCAGGCAAACTTCTTAAGGAGATGTAAGAGGATATCT 4980  
 Qy 4981 CTTTCTCTAGCTGAGAGAGAGTGTCTTAATTAATTAATTAATTAATTAATTTCCCT 5040  
 Db 4981 CTTTCTCTAGCTGAGAGAGAGTGTCTTAATTAATTAATTAATTAATTAATTTCCCT 5040  
 Qy 5041 GTCTTGTATTTAGATTTGTACCAACAGGCGGCTTGGCTGAAAGGAACTGAAGG 5100  
 Db 5041 GTCTTGTATTTAGATTTGTACCAACAGGCGGCTTGGCTGAAAGGAACTGAAGG 5100  
 Qy 5101 CCGGAGGAGGAGAAATAGATGAAACAAACAAACAAACAAACCTTCCCTAAGCTCT 5160  
 Db 5101 CCGGAGGAGGAGAAATAGATGAAACAAACAAACAAACAAACCTTCCCTAAGCTCT 5160  
 Qy 5161 ACAAACATTTTGGCCCAAGAAATGTCAAGAAATCTCAAAATCAACAGATTCGAG 5220  
 Db 5161 ACAAACATTTTGGCCCAAGAAATGTCAAGAAATCTCAAAATCAACAGATTCGAG 5220  
 Qy 5221 TACAAGAGTGTATGTATGTAGCTGAGCAAGGCTGACATCACTCACTCACTCACTCA 5280  
 Db 5221 TACAAGAGTGTATGTATGTAGCTGAGCAAGGCTGACATCACTCACTCACTCACTCA 5280  
 Qy 5281 AAAGTCCAGGCTGCAAAATTAATCTGATGCTCAATGATGATGATGATGATGATGAT 5340  
 Db 5281 AAAGTCCAGGCTGCAAAATTAATCTGATGCTCAATGATGATGATGATGATGATGAT 5340  
 Qy 5341 CAGATCAAGATTTCAAGATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 5400  
 Db 5341 CAGATCAAGATTTCAAGATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 5400  
 Qy 5401 TACGTCAGAGATCTGAAAGGCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 5460  
 Db 5401 TACGTCAGAGATCTGAAAGGCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 5460  
 Qy 5461 GCTGTAAAGATGAGGATTTTAAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA 5520  
 Db 5461 GCTGTAAAGATGAGGATTTTAAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTA 5520  
 Qy 5521 TCAAGAAAGCTTCCGCTCTTTGCAAAATGATATGAGAGAGAGAGAGAGAGAGAGAG 5580  
 Db 5521 TCAAGAAAGCTTCCGCTCTTTGCAAAATGATATGAGAGAGAGAGAGAGAGAGAGAG 5580  
 Qy 5581 TGTCTGATGATTTGAG 5640  
 Db 5581 TGTCTGATGATTTGAG 5640  
 Qy 5641 AGGATTTGTATATTTTATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTG 5700  
 Db 5641 AGGATTTGTATATTTTATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTG 5700

RESULT 4  
 AAF25907/C  
 ID AAF25907 standard; RNA; 5676 BP.

AAF25907;  
 19-APR-2001 (first entry)  
 Human tumor suppressor gene p51 promoter associated RNA SEQ ID 4.

Tumor suppressor; p51; cell death; cell proliferation; cancer;  
 cytotoxic; gene therapy; screening; de.

Homo sapiens.

W0200100818-A1.

04-JAN-2001.

XX

XX

XX

XX

XX

XX

28-JUN-2000; 2000MO-JP004261.

29-JUN-1999; 99JP-00183195.

(NIPK) NIPPON KAYAKU KK.

(SARA) SAKAI T.

Sakai T, Kagaya S, Sato T, Sukenaga Y, Fujii H;

WPI; 2001-112452/12.

New DNA sequences encoding the human p51 promoter domain for screening  
FT for compounds that modify p51 promoter activity and for diagnosing and  
FT treating cancer.

Disclosure: Page 42-45; 60pp; Japanese.

This invention describes novel DNA sequences (1) encoding the human p51  
CC promoter domain (which may include the 5'-untranslated sequence) or  
CC derived from it by addition, deletion and/or substitution of one or more  
CC bases. The invention also describes (1) expression plasmids including (1)  
CC (2) host cells transformed by (1); (3) DNA probe binding to all or  
CC part of (1); (4) cloning (1) using (3); (5) DNA (and RNA corresponding to  
CC (1) inhibiting the expression of p51, which includes the antisense  
CC sequence of (1); (6) screening compounds for their ability to modify p51  
CC promoter activity, by observing their effect on cells transformed by (1);  
CC (7) compounds identified by (6); and (8) drug compositions containing  
CC (7). The products of the invention have cytotoxic activity and can be  
CC used for gene therapy. (1) is used to screen for compounds that modify  
CC p51 promoter activity. (1) and RNA corresponding to it are used to  
CC inhibit the expression of p51. (1) and the identified compounds are used  
CC for the diagnosis and treatment of cancer

Sequence 5676 BP; 1723 A; 1116 C; 1053 G; 0 T; 1784 U; 0 Other;

Query Match 95.2%; Score 5676; DB 5; Length 5676;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CAGCTGTCAGGAGATGTCGAGAAAGAAAGCCACCCACATGCTTCGACACTGGCTGT	60
Db	5676	CAGCTGTCAGGAGATGTCGAGAAAGAAAGCCACCCACATGCTTCGACACTGGCTGT	5617
Qy	61	GACCTTGGAGGGATCAGGTTGTCTGTAAAGAAAGTGGCAACCTCCGCGCCGAT	120
Db	5616	GACCTTGGAGGGATCAGGTTGTCTGTAAAGAAAGTGGCAACCTCCGCGCCGAT	5557
Qy	121	TGGCTCTGTTCCCTTGACATGCGCTCTTCTTGGACACTCCCTTAAGCATCTTCTG	180
Db	5556	TGGCTCTGTTCCCTTGACATGCGCTCTTCTTGGACACTCCCTTAAGCATCTTCTG	5497
Qy	181	ACATTAACTTAACATAATGTTTATTTGATGATTTCACTGACCTGAAGAGATGAG	240
Db	5496	ACATTAACTTAACATAATGTTTATTTGATGATTTCACTGACCTGAAGAGATGAG	5437
Qy	241	GTCATATGAGAAAGACATGGCTAAGTTGCAATGCACTTCTTTCATTGAATTA	300
Db	5436	GTCATATGAGAAAGACATGGCTAAGTTGCAATGCACTTCTTTCATTGAATTA	5377
Qy	301	AGTCATTGCAATACCAATTCAGTTTACTTAAGTTCTAGGCCACGCTTACTCTTAAT	360
Db	5376	AGTCATTGCAATACCAATTCAGTTTACTTAAGTTCTAGGCCACGCTTACTCTTAAT	5317
Qy	361	GTCAGACTGAGCAATATTTAGGTCMAAGTTGGAAGATTAGACAGATCCTCTCAAT	420
Db	5316	GTCAGACTGAGCAATATTTAGGTCMAAGTTGGAAGATTAGACAGATCCTCTCAAT	5257
Qy	421	CAGAACTTTGGCTTCACTTAATAATAGATTTGTGTGTCGTCGAGCTATGTA	480
Db	5256	CAGAACTTTGGCTTCACTTAATAATAGATTTGTGTGTCGTCGAGCTATGTA	5197
Qy	481	CAGAAAGTGTCTACAAATTAATAATCCAACTCAGCTCTTCAATTTAGCAATAGT	540

Db	5196	CAGAAAGTGTCTACAAATTAATAATAACCAAACTCAGCTCTCAATTTAGCAATAGT	5137
Qy	541	TGGTGAATTTATTCACACACCTCCCTCTCTGAAAGTTCTTCTGCTCCCTCACTATA	600
Db	5136	TGGTGAATTTATTCACACACCTCCCTCTCTGAAAGTTCTTCTGCTCCCTCACTATA	5077
Qy	601	AATCGAAGATGACCTGGAAGGCTTAGACCTGAGCTTCACTTACCTGACACAAAGAT	660
Db	5076	AATCGAAGATGACCTGGAAGGCTTAGACCTGAGCTTCACTTACCTGACACAAAGAT	5017
Qy	661	TCAATTTCTCTATCTCATAGTACAGAGCTCCAGAGCTCTAGGAAACATGCAATCAT	720
Db	5016	TCAATTTCTCTATCTCATAGTACAGAGCTCCAGAGCTCTAGGAAACATGCAATCAT	4957
Qy	721	CTGCTTAAGCCCTCTGCTGGGCAATCTGTTTTCACATGGCCGTGAACCTATGCTCT	780
Db	4956	CTGCTTAAGCCCTCTGCTGGGCAATCTGTTTTCACATGGCCGTGAACCTATGCTCT	4897
Qy	781	TTCCTTGTTTACAGAACCTTATTTCTTGTGAAAACCTCTGCTGATGATGTAAGG	840
Db	4896	TTCCTTGTTTACAGAACCTTATTTCTTGTGAAAACCTCTGCTGATGATGTAAGG	4837
Qy	841	CCATCAGTCACATGATCAGGCTCTCCGCGCAACCTGGACCTTCTTTTGGAAAT	900
Db	4836	CCATCAGTCACATGATCAGGCTCTCCGCGCAACCTGGACCTTCTTTTGGAAAT	4777
Qy	901	TGAATCTTAAGCTGAATTAAGTTGTAATTAATTAATTAATTAATTAATTAATTAAT	960
Db	4776	TGAATCTTAAGCTGAATTAAGTTGTAATTAATTAATTAATTAATTAATTAATTAAT	4717
Qy	961	AGTGGCTTTCAGAAAGTACCTGATTCCTTATTTCTTAAGTCCCTGAAATTAATTAAT	1020
Db	4716	AGTGGCTTTCAGAAAGTACCTGATTCCTTATTTCTTAAGTCCCTGAAATTAATTAAT	4657
Qy	1021	CCGCTTAAGCCCTCTGAGATGATGATTTTTCACATGCTCTTATTAAGTCT	1080
Db	4656	CCGCTTAAGCCCTCTGAGATGATGATTTTTCACATGCTCTTATTAAGTCT	4597
Qy	1081	GAAATTTCAATATTTCTTCTTCAATTTCAATTTGTTTGTGTTGTTGTTGTTT	1140
Db	4596	GAAATTTCAATATTTCTTCTTCAATTTCAATTTGTTTGTGTTGTTGTTGTTT	4537
Qy	1141	GCTTTTGAAGGACAGATCAGTTCTGTGTTTATCCCAAGAACTGATGATACA	1200
Db	4536	GCTTTTGAAGGACAGATCAGTTCTGTGTTTATCCCAAGAACTGATGATACA	4477
Qy	1201	TCTTTCCTTTTAAATAAATAGTATCTAAGCTGAAAGAGTGGCTACCTGAGG	1260
Db	4476	TCTTTCCTTTTAAATAAATAGTATCTAAGCTGAAAGAGTGGCTACCTGAGG	4417
Qy	1261	TCTGGAGTAAAGTTAGTACAGAGCTGTACTAACCCAGGTTAGCCAACTGCTTACAC	1320
Db	4416	TCTGGAGTAAAGTTAGTACAGAGCTGTACTAACCCAGGTTAGCCAACTGCTTACAC	4357
Qy	1321	AACATTTGCTCTCTCTGAGATTATAGAGCTCTGGAAGAAAGTACTATTTGG	1380
Db	4356	AACATTTGCTCTCTCTGAGATTATAGAGCTCTGGAAGAAAGTACTATTTGG	4297
Qy	1381	CAAGACCTCAGAGACCAAGAACAGTTCTGGAAATGATGATTAACCTTTAA	1440
Db	4296	CAAGACCTCAGAGACCAAGAACAGTTCTGGAAATGATGATTAACCTTTAA	4237
Qy	1441	AGTTGTTGACCTTCTGACATTAATGTTATCTAAGACCAATTTCAATTTAATAGCT	1500
Db	4236	AGTTGTTGACCTTCTGACATTAATGTTATCTAAGACCAATTTCAATTTAATAGCT	4177
Qy	1501	AAAGCAACCAAGAGATCAAGAGTTTACAGAGTTATATTAAGTGAATTTCTGATAC	1560
Db	4176	AAAGCAACCAAGAGATCAAGAGTTTACAGAGTTATATTAAGTGAATTTCTGATAC	4117
Qy	1561	AGTTGACATGATATATCTCAAGAACATGGAATGAATGAACCAAAAAATCCAGAC	1620
Db	4116	AGTTGACATGATATATCTCAAGAACATGGAATGAATGAACCAAAAAATCCAGAC	4057



Db 1836 CCAAGAGTGTCCAGAGCACAGACATCAAGTAACTGGAGAGCTGTGAAATGCAAG 1837  
Qy 3841 AGCTTAGGCTCAGCCCAAGACCTACTGAACAGAACTGCAATTAACAAGATTTCTAGCT 3900  
Db 1836 AGCTTAGGCTCAGCCCAAGACCTACTGAACAGAACTGCAATTAACAAGATTTCTAGCT 1777  
Qy 3901 GGCTCAGGGGACATTAACCTGAGAGCTGCACTAGAAATCTGACCTCACTTC 3960  
Db 1776 GGCTCAGGGGACATTAACCTGAGAGCTGCACTAGAAATCTGACCTCACTTC 1717  
Qy 3961 AATTAAATGGAATCACTGGGCTGTGTCAAGGAAATTTGATTTTAAATTTGGA 4020  
Db 1716 AATTAAATGGAATCACTGGGCTGTGTCAAGGAAATTTGATTTTAAATTTGGA 1657  
Qy 4021 CCTTCATTTAGGCTATCTAATTTGCTAATAGCAGGAGAAACCAACTCTTAACT 4080  
Db 1656 CCTTCATTTAGGCTATCTAATTTGCTAATAGCAGGAGAAACCAACTCTTAACT 1597  
Qy 4081 GGAATTAACAATCTAATTAATTAAGTTAAGCAATCTCCCTTAAGTTTACATTTG 4140  
Db 1596 GGAATTAACAATCTAATTAATTAAGTTAAGCAATCTCCCTTAAGTTTACATTTG 1537  
Qy 4141 TGGAGCAAGCTGTGATTTGGCTGAGGCTCAGGCGGCTGTGTTGATTTCAATTT 4200  
Db 1536 TGGAGCAAGCTGTGATTTGGCTGAGGCTCAGGCGGCTGTGTTGATTTCAATTT 1477  
Qy 4201 CACAGATGTTAGCCGCTCTGGGAGTAAAGAGAGAAATGTCAGTTTAAATAGCT 4260  
Db 1476 CACAGATGTTAGCCGCTCTGGGAGTAAAGAGAGAAATGTCAGTTTAAATAGCT 1417  
Qy 4261 TCTCCCTTCATCTGCTGCTGAGCAACAATTAATTTTATGAAAACACTTTGAGT 4320  
Db 1416 TCTCCCTTCATCTGCTGCTGAGCAACAATTAATTTTATGAAAACACTTTGAGT 1357  
Qy 4321 TAGATTTACTTACAGGAAATGTCAAATTTCTGAAAGGGCTTTAGATTTCTCAAC 4380  
Db 1356 TAGATTTACTTACAGGAAATGTCAAATTTCTGAAAGGGCTTTAGATTTCTCAAC 1297  
Qy 4381 TTTGAGATCTAGTATGTCACCTAATTTAAGAGTGTCTGTGATCTAGAGGGTGAAGGA 4440  
Db 1296 TTTGAGATCTAGTATGTCACCTAATTTAAGAGTGTCTGTGATCTAGAGGGTGAAGGA 1237  
Qy 4441 AGATGGAATCAACATGTTAGTACCGTTTATGATACAGAGTGTCTTTTCCCTG 4500  
Db 1236 AGATGGAATCAACATGTTAGTACCGTTTATGATACAGAGTGTCTTTTCCCTG 1177  
Qy 4501 TTGAGTCTATCTAAGTGAAGCTTCTGAAATCAATTTCAATTTCCAAATTCACAA 4560  
Db 1176 TTGAGTCTATCTAAGTGAAGCTTCTGAAATCAATTTCAATTTCCAAATTCACAA 1117  
Qy 4561 ACCAGATAAGTTTACAGCCCATATTCGAAAGAAATTAATTTTGTGTAGACTT 4620  
Db 1116 ACCAGATAAGTTTACAGCCCATATTCGAAAGAAATTAATTTTGTGTAGACTT 1057  
Qy 4621 TCCGATTAATTAACCTGATTTTGGAAATTAAGAAATTTATGATTTCTTTGAGTAG 4680  
Db 1056 TCCGATTAATTAACCTGATTTTGGAAATTAAGAAATTTATGATTTCTTTGAGTAG 997  
Qy 4681 GTCAAGTCAAGCAAAACCAAAACAGCAAAACCTGTAGACATTAAGAAATAGAGTAG 4740  
Db 996 GTCAAGTCAAGCAAAACCAAAACAGCAAAACCTGTAGACATTAAGAAATAGAGTAG 937  
Qy 4741 CCGAGTGAAGATTAATAAATCTAGATTTTAAACAGCAATTTGAAATTAATTT 4800  
Db 936 CCGAGTGAAGATTAATAAATCTAGATTTTAAACAGCAATTTGAAATTAATTT 877  
Qy 4801 GTGACATTCAGATTTCTACAAATTAATTAATTTTCAATTTTAAATTTTAAAGAAA 4860  
Db 876 GTGACATTCAGATTTCTACAAATTAATTAATTTTCAATTTTAAATTTTAAAGAAA 817  
Qy 4861 TTACATTAATTAATTAATTAATTAATTAATTTTCAATTTTAAATTTTAACTCAATTA 4920

Db 816 TTACTATATTAATTAATTAATTAATTAATTAATTTAAGATTAATTAATTAATTA 757  
Qy 4921 AGGTATATTTCTTTTATTTGCGGTGAGGCAAAAGCTTTAAAGGGATGTCAGAAATCT 4980  
Db 756 AGGTATATTTCTTTTATTTGCGGTGAGGCAAAAGCTTTAAAGGGATGTCAGAAATCT 697  
Qy 4981 CTTTCTCTTACGTGAGAGAAAGTGAAGTTCTAAGTTAAATTAATTAATTAATTAATTA 5040  
Db 696 CTTTCTCTTACGTGAGAGAAAGTGAAGTTCTAAGTTAAATTAATTAATTAATTAATTA 637  
Qy 5041 GTCTTGTCTATTTGAGATTTGTCACCAACAGGCGGTGTGCTGAAAGGGAACTGAAGG 5100  
Db 636 GTCTTGTCTATTTGAGATTTGTCACCAACAGGCGGTGTGCTGAAAGGGAACTGAAGG 577  
Qy 5101 CCGGAGGAGAGGAAATTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5160  
Db 576 CCGGAGGAGAGGAAATTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 517  
Qy 5161 ACAAAACATTTTAGCCCAAGAAATGTCACAAATTTCTCAATTAACATGATTCAGAG 5220  
Db 516 ACAAAACATTTTAGCCCAAGAAATGTCACAAATTTCTCAATTAACATGATTCAGAG 457  
Qy 5221 TACAAGAAAGTGTATGATGTCAGAGAGGTCGACATGATGATGATGATGATGATGATGAT 5280  
Db 456 TACAAGAAAGTGTATGATGTCAGAGAGGTCGACATGATGATGATGATGATGATGATGAT 397  
Qy 5281 AAAGTCCAGGCTGCTGAATTTAACTTGAATCCATTCATGCGAGATCCAAATCAGAGA 5340  
Db 396 AAAGTCCAGGCTGCTGAATTTAACTTGAATCCATTCATGCGAGATCCAAATCAGAGA 337  
Qy 5341 GAGATCAGAAAGTTCAGAGATGCTCAGCTCAGCTCAAAATTTGCCAACAACAGTGTGCTACTA 5400  
Db 336 GAGATCAGAAAGTTCAGAGATGCTCAGCTCAGCTCAAAATTTGCCAACAACAGTGTGCTACTA 277  
Qy 5401 TACGTCAAGACCTTCGAACCTGAGAGAGGGGAGAAACAAGTGAAGAGATGAGGCA 5460  
Db 276 TACGTCAAGACCTTCGAACCTGAGAGAGGGGAGAAACAAGTGAAGAGATGAGGCA 217  
Qy 5461 CCTGTGAAGATGAGTGTATTAAGATTTAGTCAATTTGATGATGATGATGATGATGATGAT 5520  
Db 216 CCTGTGAAGATGAGTGTATTAAGATTTAGTCAATTTGATGATGATGATGATGATGATGAT 157  
Qy 5521 TCAAGAAAGCTCCGCTCTTTGCAAAATGATGATGAGAGAGATGCTTAACCTTCTA 5580  
Db 156 TCAAGAAAGCTCCGCTCTTTGCAAAATGATGATGAGAGAGATGCTTAACCTTCTA 97  
Qy 5581 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5640  
Db 96 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 37  
Qy 5641 AGGTATTTGTGTAATTTAATTAATTTATGATTTCTCCGT 5676  
Db 36 AGGTATTTGTGTAATTTAATTTAATTTATGATTTCTCCGT 1  
  
RESULT 5  
AAFS5906/c  
ID AAF55906 standard; DNA; 5676 BP.  
XX  
XX AAF55906;  
XX  
XX 19-APR-2001 (first entry)  
XX  
XX Human tumor suppressor gene p51 promoter associated DNA SEQ ID 3.  
XX  
XX Tumor suppressor; p51; cell death; cell proliferation; cancer;  
XX  
XX cytostatic; gene therapy; screening; de.  
XX  
XX Homo sapiens.  
XX  
XX WO200100818-A1.  
XX  
XX 04-JAN-2001.  
XX  
XX PD









QY	3781	CCAAAGTGTGGTCCAAGACACAGCAGCATAAAGTAACCTGGGAACTGTGTAGAAATGAG	3844
Db	1896	CCAAAGTGTGGTCCAAGACACAGCAGCATAAAGTAACCTGGGAACTGTGTAGAAATGAG	1837
QY	3841	AGCTATAGGCGCTCAACCCAGACCTACGAAACCAACATCTGCATTAAACAAGTTCTAGCT	3900
Db	1836	AGCTATAGGCGCTCAACCCAGACCTACGAAACCAACATCTGCATTAAACAAGTTCTAGCT	1777
QY	3901	GCTCTACGGGCAATTAAACCTTGAAAGCTCTGCACTAGAAATCTTCACTCCACCTTTC	3966
Db	1776	GCTCTACGGGCAATTAAACCTTGAAAGCTCTGCACTAGAAATCTTCACTCCACCTTTC	1711
QY	3961	ATTATTAATGGAAATCACTTGGGCTGTGTCCAGGAAATTAATTTATTTAATTCAGAA	4020
Db	1716	ATTATTAATGGAAATCACTTGGGCTGTGTCCAGGAAATTAATTTATTTAATTCAGAA	1657
QY	4021	CCCTCTATTAGGCTCACTATATTTGCTAATACAGGAAAGAAAGCCAAACTCTTAAC	4088
Db	1656	CCCTCTATTAGGCTCACTATATTTGCTAATACAGGAAAGAAAGCCAAACTCTTAAC	1597
QY	4081	GCATTAACAACATCTAATTAATTAATAGTAAGCAATCTCCCTTAAGTTTACATTTTG	4140
Db	1596	GCATTAACAACATCTAATTAATTAATAGTAAGCAATCTCCCTTAAGTTTACATTTTG	1537
QY	4141	TGAGCAAGCTGTTTCAATTTGGCTGAGGCTCAAGGCTGAGTGTGAAATTCACAAAT	4200
Db	1536	TGAGCAAGCTGTTTCAATTTGGCTGAGGCTCAAGGCTGAGTGTGAAATTCACAAAT	1477
QY	4201	CACAGATGTTAGCCGCTCTCGGGCTAGATAAGTAAGCAAGATGTCAGTTTAAATAGCT	4266
Db	1476	CACAGATGTTAGCCGCTCTCGGGCTAGATAAGTAAGCAAGATGTCAGTTTAAATAGCT	1417
QY	4261	TCTCCCTTCATCTCTGGCTGAAGCAACAATTAATTTTATAGAAACAATTTGAGT	4320
Db	1416	TCTCCCTTCATCTCTGGCTGAAGCAACAATTAATTTTATAGAAACAATTTGAGT	1357
QY	4321	TAGATTAATCTTAAGAGGAATGTCAAAATTTCTCGAAAGGCTTAAAGTTGTCTCAAC	4380
Db	1356	TAGATTAATCTTAAGAGGAATGTCAAAATTTCTCGAAAGGCTTAAAGTTGTCTCAAC	1297
QY	4381	TTTGAACATCTAGATGTCACTAATTAACAGTGTGTCTGTGACTAGGAGGCTGAAGGGA	4440
Db	1296	TTTGAACATCTAGATGTCACTAATTAACAGTGTGTCTGTGACTAGGAGGCTGAAGGGA	1237
QY	4441	AGATGTAACTCAACATGTTAGTGAACGTTAGATCAACAGAGTGATTTTTTCCCCCG	4500
Db	1236	AGATGTAACTCAACATGTTAGTGAACGTTAGATCAACAGAGTGATTTTTTCCCCCG	1177
QY	4501	TTGAGGTCTATCTCACTAGCTGCTGTGAACATATCAATTCATTAATTCCAAATCCCAAA	4566
Db	1176	TTGAGGTCTATCTCACTAGCTGCTGTGAACATATCAATTCATTAATTCCAAATCCCAAA	1117
QY	4561	ACACAGATAGCTTAACAGCCCATATCAAGAAAGGAATTAATTAATTTGTGTGTAGCTT	4620
Db	1116	ACACAGATAGCTTAACAGCCCATATCAAGAAAGGAATTAATTAATTTGTGTGTAGCTT	1057
QY	4621	TCTGTATTAATCACTGATTTGGGAATATGAACAATTAATGTGTTCTCTTTCGAGTAG	4680
Db	1056	TCTGTATTAATCACTGATTTGGGAATATGAACAATTAATGTGTTCTCTTTCGAGTAG	997
QY	4681	GTCACAGTCAAAGCAAAACCAAAACACGAAAACTGTAGACATTAAGAAATACATGTGAG	4740
Db	996	GTCACAGTCAAAGCAAAACCAAAACACGAAAACTGTAGACATTAAGAAATACATGTGAG	937
QY	4741	CCGAGCTGAGGATTTAAATTAACATGAATTTTTTTTAAACGAGCATTTGAATTAATTT	4800
Db	936	CCGAGCTGAGGATTTAAATTAACATGAATTTTTTTTAAACGAGCATTTGAATTAATTT	877
QY	4801	GTCAGCTTCAGAAATCTTCAACATTAATTAATTTTCCATTTTAAATCTTTAGAGAAA	4860
Db	876	GTCAGCTTCAGAAATCTTCAACATTAATTAATTTTCCATTTTAAATCTTTAGAGAAA	817
QY	4861	TTACATATTAATATGTAAGTACATGTGATGTGAGTAGAATTTAATCACTAATTA	4920

Dd		816	TTCATATTATATATAGTACATGTGCTATGTGGTAGCATTTTAATCAATATA	757
Oy		4921	AGGTTAATTTTCTTTTATTCGGGCTCAGGCACAAGCTTCTAAAGGGATGTGAAGGATATCT	4980
Dd		756	AGGTATATTTTCTTTTATTCGGGCTCAGGCACAAGCTTCTAAAGGGATGTGAAGGATATCT	697
Oy		4981	CTTTCCTTAGCTGAGAAGAGTGAGTTCTAAGTTAATATTAATCAAGCAATTTCCCT	5040
Dd		696	CYTTCCTTAGCTGAGAAGAGTGAGTTCTAAGTTAATATTAATCAAGCAATTTCCCT	637
Oy		5041	GTTTTGCTATTGAGATTGTGACACACAACAGCGGGTGGCTGAAAAGGAAATCGAAGG	5100
Dd		636	GTTTTGCTATTGAGATTGTGACACACAACAGCGGGTGGCTGAAAAGGAAATCGAAGG	577
Oy		5101	CGGGAGGGAGGAAATATGATGAAAAAACAAACAAACTCCCTTAAGCAGCTCT	5160
Dd		576	CGGGAGGGAGGAAATATGATGAAAAAACAAACAAACTCCCTTAAGCAGCTCT	517
Oy		5161	ACAAACATTTTATGCCCCGAGAAATAGTCACGAAATCTCTTAATCAACCACTATCCAGA	5220
Dd		516	ACAAACATTTTATGCCCCGAGAAATAGTCACGAAATCTCTTAATCAACCACTATCCAGA	457
Oy		5221	TACAGGAAGTGTATGTAGCTGAGACAGGATGAGACATCATCAGCTCAGTTCAGTTACA	5280
Dd		456	TACAGGAAGTGTATGTAGCTGAGACAGGATGAGACATCATCAGCTCAGTTCAGTTACA	397
Oy		5281	AAAGTCAGGCTGCTGAATTTAACTGTAGTGCATTAATGCGAGATCCOATCACAGA	5340
Dd		396	AAAGTCAGGCTGCTGAATTTAACTGTAGTGCATTAATGCGAGATCCOATCACAGA	337
Oy		5341	GAGATCAGAACTTCAGAGATGCCCTCCAGCTCCAAATTGCCAACACAAGTGGCTACTA	5400
Dd		336	GAGATCAGAACTTCAGAGATGCCCTCCAGCTCCAAATTGCCAACACAAGTGGCTACTA	277
Oy		5401	TACGTCAAGACTCTGAAAGCCGTGAGAGAGGGGAGAAACAACAGTAGAGAGATGCCCA	5460
Dd		276	TACGTCAAGACTCTGAAAGCCGTGAGAGAGGGGAGAAACAACAGTAGAGAGATGCCCA	217
Oy		5461	GCTGTGAGAAATCGAGTGTATTGAAATTTTATGCAATTTGATGAATCTCATTGGCTAAA	5520
Dd		216	GCTGTGAGAAATCGAGTGTATTGAAATTTTATGCAATTTGATGAATCTCATTGGCTAAA	157
Oy		5521	TCAAAGAAACGCTCCGGGCTCTTGCAGAAATATGTATGAAAGAGAGAGATGCCCAACTCTA	5580
Dd		156	TCAAAGAAACGCTCCGGGCTCTTGCAGAAATATGTATGAAAGAGAGAGATGCCCAACTCTA	97
Oy		5581	TGTCGATAGACATTTGACACCCTATTTGCTTTAAGCTCCGGGCTTTATATCTATATATAC	5640
Dd		96	TGTCGATAGACATTTGACACCCTATTTAAGCTCCGGGCTTTATATCTATATATATAC	37
Oy		5641	AGGTAATTTGTATATATTTATATATATGTTCTCCGT	5676
Dd		36	AGGTAATTTGTATATATTTATATATATGTTCTCCGT	1
<hr/>				
RESULT 6				
ABX5655/c				
ID	ABX5655 standard; DNA; 118384 BP.			
XX AC	ABX5655:			
XX DT	20-FEB-2003 (first entry)			
DE	Human autoimmune disease related protein PARI gene region #1.			
KM	PARI; human; autoimmune disease; psoriasis; type I diabetes;			
OS	rheumatoid arthritis; cation-chloride transport; gene therapy; gene; ds.			
XX	Homo sapiens.			
XX	WQ200280842-A2.			









```

FT      exon
FT      4165..4602
FT      /tag= g
FT      intron
FT      4603..31231
FT      /tag= h
FT      variation
FT      replace(4986,T)
FT      /tag= i
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(5268,C)
FT      /tag= j
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(7344,C)
FT      /tag= k
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(7414,C)
FT      /tag= l
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(8113,A)
FT      /tag= m
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(8334,T)
FT      /tag= n
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(9233,A)
FT      /tag= o
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(9255,T)
FT      /tag= p
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(9746..9747,A)
FT      /tag= q
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(9747,M)
FT      /tag= r
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(11634..11635,C)
FT      /tag= s
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(11635,A)
FT      /tag= t
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(13721,T)
FT      /tag= u
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(13884,C)
FT      /tag= v
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(13987,A)
FT      /tag= w
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(14376..14377,C)
FT      /tag= x
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(15045,T)
FT      /tag= y
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(15742,G)
FT      /tag= z
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(17482,G)
FT      /tag= aa
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(17688,G)
FT      /tag= ab
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(17992,C)
FT      /tag= ac
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(19132..19133,G)
FT      /tag= ad
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(20105,G)
FT      /tag= ae
FT      /standard_name= "Single nucleotide polymorphism"

```

```

FT      variation
FT      replace(24194..24195,T)
FT      /tag= af
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(24219..24220,C)
FT      /tag= ag
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(24220,A)
FT      /tag= ah
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(24278,A)
FT      /tag= ai
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(30132,C)
FT      /tag= aj
FT      /standard_name= "Single nucleotide polymorphism"
FT      exon
FT      31232..31506
FT      /tag= ak
FT      31507..33591
FT      /tag= al
FT      33592..33928
FT      /tag= am
FT      variation
FT      replace(33800,M)
FT      /tag= an
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(34513,A)
FT      /tag= ao
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(34821,G)
FT      /tag= ap
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(34983,A)
FT      /tag= aq
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(35153,T)
FT      /tag= ar
FT      /standard_name= "Single nucleotide polymorphism"
FT      variation
FT      replace(36368,T)
FT      /tag= as
FT      /standard_name= "Single nucleotide polymorphism"
FT      PN
FT      WO200192496-A2.
FT      XX
FT      PD
FT      06-DEC-2001.
FT      XX
FT      31-MAY-2001; 2001WO-US017510.
FT      XX
FT      01-JUN-2000; 2000US-0208331P.
FT      PR
FT      18-DEC-2000; 2000US-00738894.
FT      XX
FT      (APPL-) APPLEBA CORP.
FT      PA
FT      Gugler K, DI Francesco V, Beasley EM;
FT      PI
FT      WPI: 2002-130533/17.
FT      DR
FT      P-PDS; AMR17136.
FT      DR
FT      XX
FT      New isolated human kinase proteins and nucleic acids, useful as a major
FT      PT
FT      target for drug action and development, particularly for screening
FT      modulators of the kinase peptides.
FT      XX
FT      Claim 4; Fig 3; 87pp; English.
FT      XX
FT      The invention relates to human kinase proteins that are related to G-
FT      CC
FT      protein coupled receptor kinase subfamily. Human kinase gene is located
FT      CC
FT      on chromosome 3. The kinase peptide and nucleic acid are useful in the
FT      CC
FT      development of human therapeutic and diagnostic compositions. The peptide
FT      CC
FT      is useful as a major target for drug action and development, and is
FT      CC
FT      valuable to the field of pharmaceutical development to identify and
FT      CC
FT      characterise modulators of the kinase. The proteins may also be used to
FT      CC
FT      raise antibodies or to elicit an immune response, as a reagent in assays
FT      CC
FT      designed to quantitatively determine levels of the protein in biological
FT      CC
FT      fluids and as markers for tissues in which the corresponding protein is
FT      CC
FT      preferentially expressed (either constitutively or at a particular stage

```





FT	/tag= b
FT	/strand_name= "Single nucleotide polymorphism"
FT	3000. .21730
FT	/tag= c
FT	/product= "human secreted protein"
FT	/note= "contains 2 introns"
FT	3000. .3087
FT	/tag= d
FT	/number= 1
FT	3088. .7777
FT	/tag= e
FT	/number= 1
FT	replace(5184. .5185, C)
FT	/tag= f
FT	/strand_name= "Single nucleotide polymorphism"
FT	7778. .7909
FT	/tag= g
FT	/number= 2
FT	7910. .21575
FT	/tag= h
FT	/number= 2
FT	replace(8753, G)
FT	/tag= i
FT	/strand_name= "Single nucleotide polymorphism"
FT	replace(14860, G)
FT	/tag= j
FT	/strand_name= "Single nucleotide polymorphism"
FT	replace(15865, A)
FT	/tag= k
FT	/strand_name= "Single nucleotide polymorphism"
FT	replace(16060, G)
FT	/tag= l
FT	/strand_name= "Single nucleotide polymorphism"
FT	replace(16510, C)
FT	/tag= m
FT	/strand_name= "Single nucleotide polymorphism"
FT	replace(19213, T)
FT	/tag= n
FT	/strand_name= "Single nucleotide polymorphism"
FT	replace(20356, C)
FT	/tag= o
FT	/strand_name= "Single nucleotide polymorphism"
FT	21576. .21727
FT	/tag= p
FT	/number= 3
FT	replace(21904, G)
FT	/tag= q
FT	/strand_name= "Single nucleotide polymorphism"
FT	replace(22425, G)
FT	/tag= r
FT	/strand_name= "Single nucleotide polymorphism"
FT	replace(22575, G)
FT	/tag= s
FT	/strand_name= "Single nucleotide polymorphism"
PN	MO2003006481-A2.
PN	23-JAN-2003.
PD	10-JUL-2002; 2002MO-USO21574.
PF	10-JUL-2002; 2003US-00901152.
PR	10-JUL-2003; 2003US-00901152.
PA	(ABPL-) APPLERA CORP.
PI	Hu S, Ketchum K, Ladunga I;
PI	WPI: 2003-210419/20.
DR	P-PSDB: ABBY1736.
XX	New isolated peptide, useful for treating a disease or condition mediated
XX	by a human secreted protein e.g. cancer.
XX	

PS Claim 1 b; Fig 3; 93pp; English.

The invention relates to a novel human secreted protein that is related to the epidermal growth factor (EGF) superfamily. The activity of the protein of the invention may be described as cytoskeletal. The protein of the invention is useful for preparing a composition for diagnosing or treating a disease or condition mediated by human secreted proteins, including cancer. The gene encoding the protein of the invention may be found on human chromosome 22. The current sequence represents the human secreted protein genomic DNA sequence.

**SQ** Sequence 58985 BP; 13362 A; 14853 C; 15805 G; 13968 T; 0 U; 997 Other;

Query Match	Score	DB #	Length
3.04	181	8	58985

Matches 271; Conservative 0; Mismatches 105; Indels 7; Gaps 2;

3250 TAGGCTATAACCTTTTAATTTTTAAATAATAGATGAGTGTGCGCTCATGCGCTGT 3305

Db 8142 TGGGCTGAACCTTTTGTCTTTTAAGAAATGCGTGGGTCGGTGCCTCAGCTCTGT 8201

3310 AATCCCACTTTGGAGCCGGGTCCGGAGCATAGCTTGAGTCCAGCAGTTTGAGACC 3365

Db 8202 AATCCAGCACTTTGGAGACCGAGGCAAGCAGATTACCTGAGGTCGAAGAGTTTGAGACC 8261

QY 3370 AGTCAGGGCAACACAGCAGACCCATATCT-----AAAAAAAAAAAAAAAACAAAACAAT 3423

Db 8262 AGCCTGGCCAACATGGCAAAACCCCATCTCTACTGAAATACACAAATACAAAAAGT 8321

3424 TACCTGGGATGGTGTGCTCACCCTGTAGTCCAGCTACACGAGAGCTGAGGCAGAGG 3481

Db 8322 TAGCCGAGCGTGGTGGCAGTCCCTATATCTCAGTACTCGTGAAGCTGAGGCTAGGAGA 8321

3484 ATCACTGAGCCAGAGGTGAGGCTGCAATGATCAATGAACTCCCTGTACACT - CAG 3544

Db 8382 ATCACTTGACCCAGAGCAGAGGTTCAGTGAAGCCAGAAAGCCACTGCATTCAG 841

3543 TCTGGGACAGTCCAGGAGCTGCTCAAGGGAATATAAGAAATAGAAATAGGAAATAGC 5800

[illegible][illegible][illegible]

## RESULT 14

ID ABZ22654 standard; DNA; 143601 BP.

AC ABZ22654;

DT 31-MAR-2003 (first entry)

DB Human epidermal growth factor related secreted protein DNA SEQ ID NO:3.

Human, secreted protein; epidermal growth factor protein subfamily;

biological activity; immune response; chromosome 22; SNP; gene;

3 XX

[illegible]

```
FT      variation      replace(1277,A)
FT      /*t00= A
```

FT	CDs	standard name= "single nucleotide polymorphism (SNP)"
FT	2076	140828

```

FT      /*tag= b
FT      /product= "EGF protein subfamily related secreted

```

protein"  
ET  
ET  
/rotes="contains introns"

FT	exon	LOC#	CONTAINS	LOCUS
F1		2076.	.2163	

```
FT      /*tag= c  
FT      /number= 1  
FT      2164. .6058  
FT      /*tag= d  
FT      /number= 1  
FT      replace(4260. .4261,C)  
FT      /*tag= e  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(4814,C)  
FT      /*tag= f  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      6059. .6190  
FT      /*tag= g  
FT      /number= 2  
FT      6191. .25226  
FT      /*tag= h  
FT      /number= 2  
FT      replace(7034,A)  
FT      /*tag= i  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(8334. .8335,G)  
FT      /*tag= j  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(13141,G)  
FT      /*tag= k  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(14146,G)  
FT      /*tag= l  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(14341,G)  
FT      /*tag= m  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(14791,C)  
FT      /*tag= n  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      15626. .115742  
FT      /*tag= et  
FT      /number= 9  
FT      replace(17494,T)  
FT      /*tag= o  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(18678,C)  
FT      /*tag= p  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(20187,A)  
FT      /*tag= q  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(20708,G)  
FT      /*tag= r  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(20858,G)  
FT      /*tag= s  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(21858,A)  
FT      /*tag= t  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(24101,A)  
FT      /*tag= u  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(24234,T)  
FT      /*tag= v  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(24974,A)  
FT      /*tag= w  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      25227. .25355  
FT      /*tag= x  
FT      /number= 3  
FT      25356. .54132  
FT      /*tag= y  
FT      /number= 3  
FT      replace(25436,A)  
FT      /*tag= z
```

```
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(25450,C)  
FT      /*tag= aa  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(25733. .24734,T)  
FT      /*tag= ab  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(26583,C)  
FT      /*tag= ac  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(26966,A)  
FT      /*tag= ad  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(26980,G)  
FT      /*tag= ae  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(26984,C)  
FT      /*tag= af  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(27044,A)  
FT      /*tag= ag  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(28933,G)  
FT      /*tag= ah  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(30219,C)  
FT      /*tag= ai  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(30949,G)  
FT      /*tag= aj  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(32377,C)  
FT      /*tag= ak  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(32395,G)  
FT      /*tag= al  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(34305,A)  
FT      /*tag= am  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(34538,R)  
FT      /*tag= an  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(35369,G)  
FT      /*tag= ao  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(35620,T)  
FT      /*tag= ap  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(40078,C)  
FT      /*tag= aq  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(42292,C)  
FT      /*tag= ar  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(43765,A)  
FT      /*tag= as  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(44837,A)  
FT      /*tag= at  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(44995. .44996,A)  
FT      /*tag= au  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(45363,T)  
FT      /*tag= av  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(46514,Y)  
FT      /*tag= ax  
FT      /standard_name= "single nucleotide polymorphism (SNP)"  
FT      replace(52639,C)  
FT      /*tag= ay  
FT      /standard_name= "single nucleotide polymorphism (SNP)"
```













ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homindaes; Homo.	
REFERENCE	1 (bases 1 to 277)
AUTHORS	Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Leverthal, C., Thornton, M., Ramchandran, R., Whittington, J., Kerner, L., Costanzo, D., McElligott, K., Booser, S., Mayr, R., Smith, E., Veloso, N., Khilav, A., Hees, J., Cochren, K., Lo, K., Offenbacher, J., Danzig, J. and Ducar, M.
TITLE	Creation of genome-wide protein expression libraries using random activation of gene expression
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE	21227151
PUBMED	11329013
COMMENT	Contact: Scott T. Cain Athersys, Inc. 3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900 Fax: 216 361 9596 Email: scotinaethersys.com High quality sequence stop: 227. Location/Qualifiers 1..277 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /cell_line="HT1080" /clone_id="Athersys RAGE Library" /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression' Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
ORIGIN	
Query Match	4.4%; Score 262; DB 12; Length 277;
Best Local Similarity	98.1%; Pred. No. 5; 9e-28;
Match 265:	Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy	3117 TCCTTTCATTAATTACTTAATTTAAATTTGTCAGATTAAATTTCTCTCTTAAGCTT 3176
Dd	270 TTTTTCATAAATTAACCTAATTTAAATTTGTGCACATTAATTTCTCTCTTAAGCTT 211
Oy	3177 AGTAACTTAGAGATTTTAAATTAACAATCAATGAATCATACATPAGCTTAATGATAT 3236
Dd	210 AGTAACTTAGAGATTTTAAATTAACAATCAATGAATCATACATPAGCTTAATGATAT 151
Oy	3237 TATTAAATACGTTAGCGCATATAAACCTTTAAATTTTTTAAAAAAAATGATAGTGSTGT 3296
Dd	150 TATTAAATACGTTAGCGCATATAAACCTTTTAAATTTTTTAAAAAAAATGATAGTGSTGT 91
Oy	3297 GGCTCAGCCCTGTAATCCCAACACTTTGGAGACCCTGGGTGGAGATACCTTGAATCCA 3356
Dd	90 GGCTCAGCCCTGTAATCCCAACACTTTGGAGACCCTGGGTGGAGATACCTTGAATCCA 31
Oy	3357 GCAGTTGAGAACCACTGACGGCAACACAGC 3386
Dd	30 GCAGTTGAGAACCACTGACGGCAACACAGC 1
RESULT 5	
LOCUS	B1497128 567 bp mRNA linear EST 28-AUG-2001
DEFINITION	d4132c05.v1 Morton Fetal Cochlea Homo sapiens cDNA clone
ACCESSION	U1497128 IMAGE:2537461 5', mRNA sequence.
VERSION	B1497128 GI:15364472
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.

REFERENCE 1 (bases 1 to 567)  
 AUTHORS Robertson,N.G., Kleierpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R. and Morton,C.C.  
 TITLE Isolation of novel and known genes from a human fetal cochlear cDNA library using subtractive hybridization and differential screening  
 JOURNAL Genome 23, 42-50 (1994)  
 MEDLINE 9513011  
 PUBMED 7829101  
 COMMENT Contact: Morton, C. C.  
 Department of Pathology and Obstetrics, Gynecology and Reproductive Biology  
 Brigham and Women's Hospital  
 75 Francis Street, Harvard Medical School, Boston, MA 02115, USA  
 Tel: 617 732 7880  
 Fax: 617 736 6396  
 Email: ccmorton@rics.bwh.harvard.edu  
 DNA sequencing and analyses were performed by National Institutes of Health Intramural Sequencing Center (NISC; see <http://www.nisc.nih.gov>).  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 Plate: LLM6320 row: F column: 10  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1..567  
 /organism="Homo sapiens"  
 /mol\_type="RNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2537481"  
 /issue\_type="cochlea"  
 /dev\_stage="16-22 week fetus"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 /clone\_lib="Morton Fetal Cochlea"  
 /note="Organ: ear; Vector: pBluescript SK-; Site: 1: EcoRI; Site: 2: XhoI; Reference: Genom 23, 42-50 (1994) Cloned unidirectionally. Primer: Oligo dt. Fetal cochlea, normal. 376 of insert <0.5 kb, 566 0.5-1.0 kb, 78 >1 kb. Uni-2AP XR Vector. Library constructed by N. Robertson, C. Morton. -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

ORIGIN

Query Match 3.1%; Score 184.6; DB 12; Length 567;  
 Best Local Similarity 68.6%; Pred. No. 4.8e-17;  
 Matches 300; Conservative 0; Mismatches 129; Indels 8; Gaps 3;

QY 3170 TTAGCTTACTTACTTGGATTTTAAATTAACAATGAAATCAATGACATGCTTAA 3229  
 DB 59 TTGCTTAAAGTTTTCATATTTGAAAAAACTCCACATCAGTACATGATCTCC 118  
 QY 3230 ATGATATTATTAAATAGCTTAAACCTTTAAATTTTAAAAAAATGATGA 3289  
 DB 119 ATATTTTCTTTTGTATCTTTGATTTTAAATTTTAAATTTAAATGAGCTG 175  
 QY 3290 GTGTGTGTGCTCATGCTTTATCCCAACCTTTGGAGAGCGGTGCGAGATAGCTT 3349  
 DB 176 GTGAGGTGCTTACTCTGTAAATCCCAACCTTTGAGAGGTGAGGTGATGATCTT 235  
 QY 3350 GAGTCAGCAGTTTGAACAGTCAAGGCAACAGAGCAAGCCCATATCTAAAAACA 3409  
 DB 216 GAGCTCAGAGTTTCAAGACCCCTGGGCAACAGAGCAAGTACTTCTCAAAAAA 295  
 QY 3410 AACCAAAACAAATTAATCTGGGTATGTGTGTCTCACTGATGCAAGTACACAGAA 3469  
 DB 296 ATAGA-----AAATTAAGTGGGCAATGTGTGTCTGATGCTTCAAGTCTGAG 351  
 QY 3470 GCTAAGCAGAGATCACTTGAAGCCAGAGTTTGAAGTGTGAGTATCATGAACGG 3529  
 DB 352 GCTAAGGCGGAGATCACTTGAAGCCAGAGTTTGAAGTGTGAGTATCATGAAGT 411  
 QY 3530 CTGCTACACT-CAGTCTGGGTGACAGTCAAGAGCTGTCTCAAAAAATATATATATA 3588  
 DB 412 CTGCTGACCTCAGCGTGGGTGACAGAGCAAGACCTGTCTCAAAAAACCCCAAAAA 471

QY 3589 AAATTAACCTTTAAAAA 3605  
 DB 472 AACGAAATATCAATAA 468

RESULT 7  
 A0275866/c 664 bp DNA linear GSS 22-NOV-1998  
 LOCUS CITBI-El-2517A7.TF CITBI-El Homo sapiens genomic clone 2517A7,  
 DEFINITION genomic survey sequence.  
 ACCESSION A0275866  
 VERSION A0275866.1 GI:3901682  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Buxarjota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 664)  
 Adams,M.D., Rounsley,S.D., Zhao,S., Base,S., Linher,K., Golden,K.,  
 Berry,K., Granger,D., Suh,F., Wible,C., Shizuya,H., Simon,M. and  
 Venter,J.C.  
 TITLE Use of a random human BAC End Sequence Database for Sequence-ready  
 Map Building  
 JOURNAL Unpublished (1998)  
 COMMENT Other\_GSSs: CITBI-El-2517A7.TF  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: madams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 End search page:  
[http://www.tigr.org/tcd/hunguen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tcd/hunguen/bac_end_search/bac_end_search.html).  
 Seq primer: M13-21  
 Class: BAC ends.  
 Location/Qualifiers  
 1..664  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="2517A7"  
 /sex="male"  
 /cell\_type="sperm"  
 /clone\_lib="CITBI-El"  
 /note="Vector: pBluescript II; Site: 1: EcoRI; Site: 2: RcoRI;  
 Caltech Human BAC Library D"

ORIGIN

Query Match 3.0%; Score 181.6; DB 28; Length 664;  
 Best Local Similarity 72.2%; Pred. No. 1.2e-16;  
 Matches 249; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

QY 3283 TAGATGATGTGTGTGCTCATGCTTTATCCCAACCTTTGGAGAGCGGTGAGAG 3342  
 DB 639 TGGCTGAGTGTGTGTGCTCACACCTGTATCCACACCTTTGGAGAGGTGAGAG 580  
 QY 3343 ATAGCTGAGTCCAGAGTTTGAACAGTCAAGGCAACAGAGCAAGCCCATATCTAA 3402  
 DB 579 ACAGTTTGAAGCCAGAGTTTGAACAGTCAAGGCAACAGAGCAAGCCCATATCT 520  
 QY 3403 AAAAAACAAACAAATTAATCTGGGTATGTGTGTCTCACTGATGCAAGTAC 3462  
 DB 519 AAAAAACAAATTAATTAAGGATTAAGGAGTGTGAGAGCAAGCCCATATCTCA 460  
 QY 3463 ACAGAGCTGAGCAGAGATCACTTGAAGCCAGAGTTTGAAGTGTGAGTATCAT 3522  
 DB 459 TCAGAGCTGAGCAGAGATCACTTGAAGCCAGAGTTTGAAGTGTGAGTATCAT 400  
 QY 3523 GAACGCGTGTACTACT-CAGTCTGGGTGACAGTCAAGAGCTGTCTCAAAAAATATA 3581

Page 6

	Db	261	CAGTTGGGAGGCCAAGGTGTGGTGCATCTTGACATCAGCAAGTTCCAGACCAAGCTTCGCG	320
	Qy	3378	CAACACGGAACAACCCTATTCTTTAAAAAAACAAAACAAAATAACTGCGGTATGGT	3437
	Qy	3318	CACATTGGAGAACCCGGGTCCGGAGCATTAACCTTGATCTCCAGCATGTTTAGACACAGTCAGGG	3377
	Db	321	CACATGGGAAAACCTCATCTTCTATCTTAAAAAAACAAAAAAA--TAGCTGGGTGTGGT	378
	Qy	3438	TGTCCTACCTTGATCCAGCTTACACAGAGAGCTGAGCGCAGAAGATCATCTTAGGCCCA	3497
	Db	379	TGCACACACCTGTATATCCAGATATCTTGGAGAGCTGAGGACGAGGATCATCTTAATCTTA	438
	Qy	3498	GGAAGTTGAGAGCTGAGAGATTCATCAGACGCGCTGCTCACT-CAGTCTGGGTGACAGTG	3556
	Db	499	GAGGCGAGAGGTTTACATGATGAGCCAAAGATTGACCACTGACCTTCAGCTTGGTGAACAG	498
	Qy	3557	CAGAGAGCTGTCTCAAATAATTAATAATAATAATAATACTTTAAAAACAAATAATTA	3616
	Db	499	AGCACACCTGCTCAAAAAAAAAATATATATATATATATATATATATATATATATATATAT	558
	Qy	3617	TTAATTTTAAAAACAC 3653	
	Db	559	TACAAATTTTATGATAC 575	
RESULT 9				
LOCUS	HZ601705	828 bp	DNA	linear GSS 08-JUN-2003
DEFINITION	MHA8P61R Human MCF7 breast cancer cell line library (MCF7_1) Homo sapiens genomic clone MCF7_1-8K2, genomic survey sequence.			
ACCESSION	HZ601705			
VERSION	HZ601705.1	GI:31510167		
KEYWORDS	GSS.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			
AUTHORS	Mumukota; Metaxoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 828)			
REFERENCE	Volk,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q., Gray,J.W. and Collins,C. End-sequence profiling: Sequence-based analysis of aberrant genomes Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7656-7701 (2003)			
JOURNAL MEDLINE PubMed	12709111 12788976			
COMMENT	Contact: Volk SV Colin Collins' lab UCSF Comprehensive Cancer Center UCSF Box 0808, San Francisco, CA 94143-0808, USA Tel: 415 502 7066 Fax: 415 502 5665 Email: volk@csc.ucsf.edu This clone is available from Amplicon Express <a href="http://www.genomex.com">http://www.genomex.com</a> Class: BNC ends. Location/Qualifiers			
FEATURES	source	1..828		
	/organism="Homo sapiens"			
	/mol_type="Genomic DNA"			
	/db_xref="taxon:9606"			
	/clone="MCF7_1-8K2"			
	/sex="female"			
	/clone_lib="Human MCF7 breast cancer cell line library (MCF7_1)"			
	/note="Vector: pECBAC1, Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express ( <a href="http://www.genomex.com">http://www.genomex.com</a> ) using their standard procedure."			
ORIGIN				
Query Match	3.0%; Score 180;	DB 28;	Length 828;	
Best Local Similarity	67.7%; Pred. NO. 1,7e-16;			







## FEATURES

Location/Qualifiers  
1. 637

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6247825"  
/issue\_type="ductal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_110"  
/note="Organ: pancreas; Vector: pOT87; Site 1: XhoI;  
Site 2: EcoRI; CDNA made by oligo-dT priming;  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the Laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 3.0%; Score 179; DB 13; Length 637;  
Best Local Similarity 75.4%; Pred. No. 2.8e-16;  
Matches 251; Conservative 0; Mismatches 75; Indels 7; Gaps 2;

QY 3285 GATGAGTGTGCTGCTCATGCTGTAATCCACACTTTGGGAAGCCGGTGGAGCAT 3344  
DB 311 GCTGGGTCACGTAACATCTGATGCTCCAGCACTTGGAGGCGGCGGACAT 370  
QY 3345 AGCTTAGTCCAGCATTTGAGACCACTGACGGCGACACAGCAAGCCCATATCTAAAA 3404  
DB 371 CACTTAGGCCACAGAGTTGAGACCAAGCTGCGGCAACAGCCAAACCTTCTCT---- 426  
QY 3405 AAACAAACAAACAAATTAACCTGGGTAAGTGTGCTCAACCTGATGTCAGTACAC 3464  
DB 427 --ACCAAAATATGAAATTTAGCTGGCAATGATGTGTGCTGATGCTCCAGCTCTC 484  
QY 3465 AGCAAGCTGACGACAGATGATCTTGAGCCCGACGAGTGTGAGCTGCAAGTATCCATCA 3524  
DB 485 AGAAGGCTGAGGCGAGATGAGCTTGAGCCCGACGAGTGTGAGCTGCAAGTATCCATCA 544  
QY 3525 ACCGCTGCTACA-CTCACTGTGGGTGACAGTCCAGACCTGTCTCAAAAATTAAT 3583  
DB 545 TTGGCGCACTGACCCGACCTGGGCAACAGAGCAACACCTGTCTCAAAAATTAAT 604  
QY 3584 AAATTAATTAATCTTTAAAAACAAATTTAA 3616  
DB 605 AAATTAATTAATCTTTAAAAACAAATTTAA 637

RESULT 14  
A1362915/c 440 bp mRNA linear EST 06-JAN-1999  
LOCUS QY8106.x1 NCI CGAP Brn25 Homo sapiens CDNA clone IMAGE:2018410 3'  
DEFINITION similar to contains Alu repetitive element; mRNA sequence.  
ACCESSION A1362915  
VERSION A1362915.1 GI:4114536  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 440)  
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BRGAP), Tumor Gene Index  
Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lannon, Ph.D.

## FEATURES

source  
1. 440

DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/URL at:  
[www-bio.1lnl.gov/dbp/image/image.html](http://www-bio.1lnl.gov/dbp/image/image.html)  
Seq primer: -400P from G1bco.  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2018410"  
/issue\_type="anaplastic oligodendroglioma"  
/lab\_host="DH10B"  
/clone\_id="NCI CGAP Brn25"  
/note="Organ: Brain; Vector: pT73D-Pac (Pharmacia) with a  
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer (5'  
TGTACCACTGTAAGTGGAGCGGCGCATGAGTTTATTTTATTTTATTTT  
T 3'); double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT73 vector.  
Library is normalized, and was constructed by Bento  
Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 3.0%; Score 178.8; DB 9; Length 440;  
Best Local Similarity 78.4%; Pred. No. 3.8e-16;  
Matches 240; Conservative 0; Mismatches 62; Indels 4; Gaps 2;

QY 3290 GTGTGTGGCTGCTATGCTGTAATCCCACTTTGGAGCGGCGGAGTACCTT 3349  
DB 304 GTGTGTGGCTGCTATGCTGTAATCCCACTTTGGAGCGGCGGAGTACCTT 245  
QY 3350 GATGTCGACGATTGAGACCACTGACGGCGACACAGCAAGCCCATATCTAAAAACA 3409  
DB 244 GAGCCGAGATTCACACCACTGAGGCAACATGCAAGACTCTCTCTACAAA--A 188  
QY 3410 AAACAAACAAATTAACCTGGTATGCTGTGCTCAACCTGATGTCAGTACACAGAA 3469  
DB 187 AAATTAATTAATCTTTAAAAACAAATTTAA 128  
QY 3470 GCTGAGCAGAGGATGATCTGAGCCAGACCTTGAGGCTGAGTATCCATCAACGCG 3529  
DB 127 GCTGAGTGGAGGATGATCTGAGCCAGAGATTTGAGGCTGATCACTATATTTGCA 68  
QY 3530 CTGCTACACT-CACTGTGGGTGACAGTCCAGACCTGTCTCAAAAATTAATTAAT 3588  
DB 67 CCATTGCACTCTGCTGCGGAGACAGATGACATCTGTCTCAAAAATTAATTAAT 8  
QY 3589 AAAATA 3594  
DB 7 AGAAAA 2

RESULT 15  
AG073447/c 674 bp DNA linear GSS 03-NOV-2001  
LOCUS AG073447  
DEFINITION Pan troglodytes DNA, clone: PTB-065D24.F, genomic survey sequence.  
ACCESSION AG073447  
VERSION AG073447.1 GI:16625249  
KEYWORDS GSS.  
SOURCE Pan troglodytes (Chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE  
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Toto, Y., Watanabe, H., and Sakaki, Y.  
BAC end sequences of library PTB  
Unpublished  
2 (bases 1 to 674)  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Toto, Y., Watanabe, H., and Sakaki, Y.

TITLE  
JOURNAL

Direct Submission  
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimpanzee@sec.riken.go.jp, URL: http://hsp.sec.riken.go.jp/,  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)

## COMMENT

Clones are derived from the chimpanzee BAC library PTB This BAC end  
was generated during the R&D process and may have higher chance of  
clone tracking errors.

## PRIMERS

Sequencing: -21M13  
LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1..674

/organism="Pan troglodytes"

/mol type="genomic DNA"

/db xref="taxon:9598"

/clone="PTB-065D24.F"

/sex="male"

/cell\_type="lymphoblast"

/clone\_id="PTB Chimpanzee Male BAC library"

## ORIGIN

## Query Match

Best Local Similarity 3.0%; Score 178.2; DB 29; Length 674;

Matches 233; Conservative 0; Mismatches 73; Indels 1; Gaps 1;

```

QY 3289 AGTGTGGTGGCTCATGCTGTAAATCCCAACACTTTGGAGCCGGGTGGAGGATAGCT 3348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 336 AGTGAAGTGGCTCATGCTGTAAATCCCAACACTTTGGAGCCGGGTGGAGGATAGCT 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3349 TGAATCCAGCGATTGAGACCACTGAGCGCAACAGCAAGCCCATATCTAATAAATAC 3408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 336 TGAATCCAGCGATTGAGACCACTGAGCGCAACAGCAAGCCCATATCTAATAAATAC 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3409 AAAACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 277 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3469 AGCTAGGCGAAGATCACTTGAAGCCGAGAGGTTGAGGCTGCAAGTATCCATGAACGC 3528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 217 GCTGAGGTGGAGATCACTTGAAGCCGAGAGGTTGAGGCTGCAAGTATCCATGAACGC 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3529 GCTGTAAGTCACTGAGTGGAGTGAAGTGAAGAGCTGTTCAATAAATAAATAAATAA 3588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 157 GCGACTGCACTCAATCTGGGTGATAGAGTGAAGCCCTGCTCAAAAAAATTAATAAAGA 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3589 AAAATAA 3595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 97 ACTATGA 91
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: September 27, 2004, 18:27:21  
Job time : 13719 secs